Chapter 22

Compactness

Introduction

We are interested in studying continuous functions. In Book D, Chapter 13, just after Example 2.17, we noted that the Intermediate Value, Boundedness and Extreme Value Theorems each hold for all continuous functions on all $d^{(1)}$-closed and $d^{(1)}$-bounded intervals in $\mathbb{R}$. This suggests that they are the result of some kind of interaction between continuity and closed and bounded intervals. In Chapter 21, we saw that the Intermediate Value Theorem can be generalised to apply to any connected metric space. In this chapter, we generalise the idea of 'closed and bounded'-ness to metric spaces, enabling us to generalise the Extreme Value Theorem.

Recall that connectedness, the characteristic of a space which allows the Intermediate Value Theorem to hold, gives an intuitive idea as to whether or not a metric space is in one 'piece'. The characteristic which enables the Extreme Value Theorem to hold turns out to be *(sequential) compactness*, which gives us another idea of 'size' or 'stretchability' for a space.

We stated, but did not prove, the Boundedness Theorem and the Extreme Value Theorem in Chapter 13.

Boundedness Theorem

Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function on $[a,b]$. Then f is bounded on $[a,b]$. That is, there is a real number M such that $|f(x)| \leq M$ for each $x \in [a,b]$.

Here 'continuous' means $(d^{(1)}_{[a,b]}, d^{(1)})$-continuous.

Extreme Value Theorem

Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function on $[a,b]$. Then there are numbers $c, d \in [a,b]$ such that

$$f(c) \leq f(x) \leq f(d) \quad \text{for each } x \in [a,b].$$

Informally, the Extreme Value Theorem tells us that a $(d^{(1)}_{[a,b]}, d^{(1)})$-continuous function on a $d^{(1)}$-closed and bounded interval of the real line is bounded and attains its bounds. Formally, this theorem tells us that two numbers, n and m, exist, with $n \leq f(x) \leq m$ for each $x \in [a,b]$, (the function is bounded) and that two numbers c and d exist, in $[a,b]$, with $f(c) = n$ and $f(d) = m$ (the function attains it bounds).

Note that we can have $n = m$ if f is a constant function.

In the first section of this chapter, we prove this result for the interval $[a,b]$, and pull out from the proof the key characteristic of $d^{(1)}$-closed and bounded intervals that is necessary for the proof to 'work'. This turns out to be the rather remarkable fact that *every* sequence in $[a,b]$ has a convergent subsequence, and it is this characteristic that gives us the definition of sequential compactness.

In the second section we formally define sequential compactness, and prove the Heine–Borel Theorem, an important result in real analysis which characterises all sequentially compact subsets of $\mathbb{R}^n$. We then, in Section 3, turn our attention to the function space $C[0,1]$ and look at the sequentially compact subsets of this space.

In Section 4 we look at compactness, a notion of size that is defined using the open set structure of a space. We go on to prove that the definitions of sequential compactness and of compactness are equivalent for all metric spaces.

Chapter guide

The ideas introduced within this chapter are subtle, and you may find that it takes some effort and practice before you are fully comfortable with them. This is normal – the material is often regarded as quite difficult.

The most important ideas of this unit appear in Sections 2 and 4. You will probably need to spend most of your time studying these two sections.

In Section 2 we introduce the idea of sequential compactness and investigate how sequences and subsequences can be used to encapsulate various properties of continuous functions on closed and bounded intervals. Later, in Section 4, we see how the ideas of an *open cover* and a *finite subcover* can be used to give an alternative approach. We show that these two approaches both lead to the same destination for a metric space in Section 5. The proof that the two notions of compactness are equivalent for metric spaces is long. You should focus on following the general ideas behind the proofs and not spend too much time studying the technical details. Section 3 looks at sequentially compact subsets of the function space $C[0,1]$ and gives a relatively simple technique that we can use to verify whether a given subset of $C[0,1]$ is sequentially compact. If you are short of time when reading this section, then you should focus on ensuring that you can use this technique.

Notation

In this chapter, when considering a subset X, of $\mathbb{R}^n$, with the metric, $d_X^{(n)}$, induced by the usual Euclidean metric on $\mathbb{R}^n$, we will often drop the subscript and use the notation $d^{(n)}$ to refer to $d_X^{(n)}$.

1 Two important theorems

In this section we prove the Boundedness Theorem and the Extreme Value Theorem for subsets of the real line (with the Euclidean metric). The Boundedness Theorem is a consequence of the Extreme Value Theorem, so we concentrate our efforts on finding a generalisation of the latter result. The Extreme Value Theorem applies to closed and bounded intervals of the real line, and we examine the proof of this result to try to pinpoint the precise aspect of 'closed and bounded'-ness upon which the proof relies.

Remark

Note that in the following discussion we use the word 'bounded' in two senses: to refer to bounded sets and to refer to bounded functions. Remember that a set $A \subseteq \mathbb{R}$ is bounded if there is a real number M such that $|x| \leq M$ for each $x \in A$; and a function $f\colon X \to \mathbb{R}$ is bounded if there is a real number M with $|f(x)| \leq M$ for each $x \in X$ (so a bounded function is a function that has a bounded image set).

See Book D, Chapter 13, Definition 2.14.

When trying to identify the precise characteristics of a mathematical object that enable us to prove a result, it is often helpful to look at what happens when we vary or drop some of the conditions. The Extreme Value Theorem applies to closed and bounded intervals of $\mathbb{R}$, so we remove each of these in turn to see if both appear necessary for the result to hold.

First we examine the effect of removing the boundedness of the domain. Consider the identity function, $\mathrm{Id}\colon [0,\infty) \to \mathbb{R}$ given by $x \mapsto x$. This (rather basic) continuous function is not bounded when defined on the unbounded interval $[0,\infty)$. This suggests that the bounded condition on the domain is a necessary one. Now consider the function $f\colon [1,2) \to \mathbb{R}$, given by $x \mapsto x^2$. The domain, $[1,2)$, is bounded but is not $d^{(1)}$-closed (when considered as a subset of $(\mathbb{R}, d^{(1)})$). The function is still bounded (as $1 \leq f(x) \leq 4$), but it does not *attain* its upper bound (as there is no value $a \in [1,2)$ with $f(a) = 4$). Therefore the domain of a real function needs to be both bounded and closed in order for the function to be bounded and to attain its bounds. So it appears reasonable to start our quest for a generalisation of the Extreme Value Theorem by examining the nature of closed and bounded subsets of $\mathbb{R}$.

Here we could, of course, have used the identity function (restricted to an appropriate interval).

1.1 Preliminaries: sequences and subsequences

We use sequences and subsequences as our main tools for investigating the nature of closed and bounded subsets of $\mathbb{R}$. Therefore we first remind ourselves of a few key definitions, and warm up with a few examples.

Let X be a set. A **sequence** in X is an unending ordered list of elements of X

$$a_1, a_2, a_3, \ldots.$$

The element a_k is the ***k*th term** of the sequence, and the whole sequence is denoted by (a_k), $(a_k)_{k=1}^{\infty}$ or $(a_k)_{k\in\mathbb{N}}$.

A real sequence (a_k) is **monotonic increasing** if for each $k \in \mathbb{N}$, $a_k \leq a_{k+1}$ and **monotonic decreasing** if for each $k \in \mathbb{N}$, $a_k \geq a_{k+1}$. It is **strictly monotonic increasing or decreasing** if the inequality is strict for each k.

The sequence $(a_{n_k})_{k=1}^{\infty}$ is a **subsequence** of the sequence (a_n) if $(n_k)_{k=1}^{\infty}$ is a strictly increasing sequence of positive integers, i.e. if $1 \leq n_1 < n_2 < n_3 < \cdots$; see Figure 1.1.

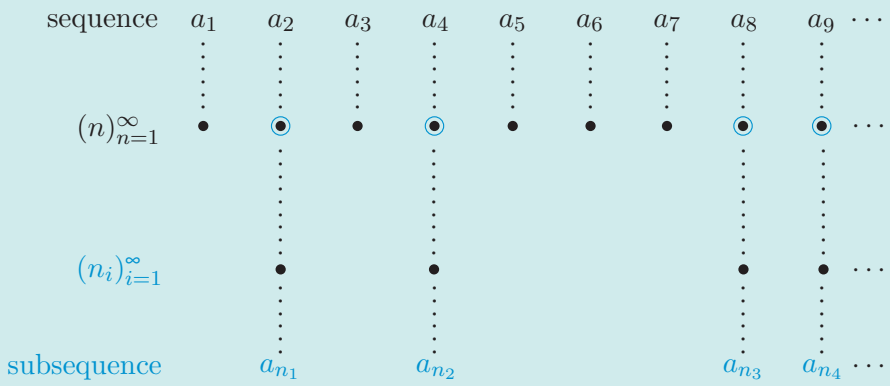


Figure 1.1

A set $A \subset \mathbb{R}$ is **bounded** if there exists a real number M such that $|x| \leq M$ for each $x \in A$.

Exercise 1.1

It is not possible to define a strictly monotonic increasing sequence for an *arbitrary* metric space by saying 'a sequence (a_n) is **strictly monotonic increasing** if for each $n \in \mathbb{N}$, $a_n < a_{n+1}$.' Why is this?

Exercise 1.2

Write down the first few terms of a sequence in $C[0, 1]$.

Exercise 1.3

See Book D, Chapter 15, Section 2.

Write down the first few terms of a sequence in Cantor space.

Exercise 1.4

See Book D, Chapter 16, Example 1.19.

Write down a null sequence in the middle-third Cantor set which does not contain the point 0 as a term.

We now focus on a concrete example and consider the set $[0, 1]$. It is very easy to write down a sequence which is not $d^{(1)}$-convergent: any alternating sequence, such as $(0, 1, 0, 1, 0, \ldots)$, will suffice. However, this sequence does have a convergent *sub*sequence. (In fact, it has infinitely many different convergent subsequences.) Since the interval $[0, 1]$ contains uncountably many points from which we can pick the terms for our sequence, and since we only need to select countably many of these to form our sequence, it may seem as if it should be easy to find a sequence that has no convergent subsequence. However, this turns out to be impossible!

See Definition 6.4 of Chapter 16.

In fact, we prove that it is impossible to select a sequence of points *from any $d^{(1)}$-closed and bounded interval* $[a, b]$ that does not have a $d^{(1)}$-convergent subsequence. In other words, we cannot pick countably many points (to form a sequence) from our uncountable set $[a, b]$, without returning arbitrarily close to some fixed point (the limit of the convergent subsequence) infinitely many times.

Of course this is not true for all metric spaces; the sequence (n) in $\mathbb{R}$ clearly has no convergent subsequence. However, we do not have to look to the whole of $\mathbb{R}$ to find an example. Consider the bounded, open interval $(0, 1)$. The sequence $\left(\frac{1}{n}\right)$ has no $d^{(1)}_{(0,1)}$-convergent subsequence. In the case of (n) in $\mathbb{R}$, we use the fact that $\mathbb{R}$ is unbounded to give us enough 'room' to find a sequence with no convergent subsequence. In the case of $\left(\frac{1}{n}\right)$ in $(0, 1)$, we use the fact that $(0, 1)$ has a closure point outside the interval in order to find 'room'. If you lived inside the interval $(0, 1)$, then you would find that the edges appeared 'fuzzy' – you could approach arbitrarily close to the points 0 or 1, but you could not reach either of them. In fact, the 'fuzzy' edges of $(0, 1)$ may look a little like the unreachable 'edges' of $\mathbb{R}$ (and you may recall that in Chapter 21 we proved that $(0, 1)$, with the Euclidean metric, was homeomorphic to $\mathbb{R}$, with the Euclidean metric).

This works because we can find a sequence of points that converges to a closure point; see Book D, Chapter 16, Worked Exercise 3.7.

Our choice of $[0, 1]$ as our closed and bounded interval was fairly random. It turns out that we can always find a convergent subsequence for any sequence in any $d^{(1)}$-closed and bounded subset of $\mathbb{R}$! As the first step on our journey towards a proof, we prove a lemma.

Lemma 1.1

Let (a_n) be a sequence of real numbers. Then (a_n) contains a monotonic subsequence – that is, there is a subsequence (a_{n_k}) such that either:

$$a_{n_1} \leq a_{n_2} \leq a_{n_3} \leq \cdots \leq a_{n_k} \leq \cdots$$

or

$$a_{n_1} \geq a_{n_2} \geq a_{n_3} \geq \cdots \geq a_{n_k} \geq \cdots .$$

The lemma does not state that the monotonic subsequence converges: the sequence $(1, -1, 2, -2, 3, -3, 4, -4, \ldots)$ contains the strictly increasing subsequence $(1, 2, 3, \ldots)$, which does not converge.

Proof Let (a_n) be a sequence in $\mathbb{R}$. We call a particular term a_i of (a_n) a *peak point* if $a_i \geq a_m$ for all $m \geq i$ (Figure 1.2). If (a_n) has infinitely many peak points, then the sequence of peak points $\left(a_{n_i}\right)$, where $n_1 < n_2, \ldots$, is a monotonic decreasing subsequence of (a_n) – and we are done.

If this isn't the case, then there are only finitely many peak points in (a_n). Let $N_1 = n_i + 1$, where a_{n_i} is the final peak point (setting $N_1 = 1$ if (a_n) has no peak points). Since a_{N_1} is not a peak point, then there must be a point a_{N_2}, further along the sequence (so $N_2 > N_1$) with $a_{N_2} > a_{N_1}$. In Exercise 1.5 we ask you to use induction to finish constructing a monotonic increasing subsequence $\left(a_{N_i}\right)$ of (a_n).

Therefore (a_n) contains a monotonic subsequence and, as (a_n) was an arbitrary sequence in $\mathbb{R}$, we have the result. ■

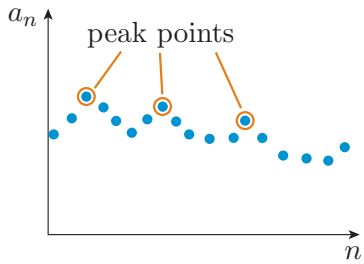


Figure 1.2

Exercise 1.5

Write a formal proof using induction to show that if there are only finitely many peak points, then we can construct a monotonic increasing subsequence (a_{N_i}) of the sequence (a_n) in the proof of Lemma 1.1.

Theorem 1.2

Let X be any $d^{(1)}$-closed and bounded subset of $\mathbb{R}$. Then every sequence in X contains a $d_X^{(1)}$-convergent subsequence.

Proof By Lemma 1.1 we know that every sequence contains a monotonic subsequence. If this is a monotonic increasing sequence then, since X is bounded above, we can apply the Monotone Convergence Theorem and conclude that this subsequence is $d^{(1)}$-convergent when considered as a sequence in $\mathbb{R}$. As X is $d^{(1)}$-closed, then the limit of this sequence must be in X. Similarly, if this is a monotonic decreasing sequence then, since X is bounded below, we can also apply the Monotone Convergence Theorem and draw the same conclusion, giving the result. ■

The Monotone Convergence Theorem is Theorem 1.22 of Chapter 13.

We now turn our attention back to the Boundedness Theorem and the Extreme Value Theorem and prove these results for $\mathbb{R}$.

Theorem 1.3 *Boundedness Theorem*

Let $f\colon [a,b] \to \mathbb{R}$ be a $(d_{[a,b]}^{(1)}, d^{(1)})$-continuous function on $[a,b]$. Then f is bounded on $[a,b]$. That is, there is a real number M such that $|f(x)| \leq M$ for each $x \in [a,b]$.

Proof The proof is by contradiction. Supposed that f is not bounded. Then for each $n \in \mathbb{N}$ we can find a point $x_n \in [a,b]$ with $|f(x_n)| > n$. This gives us two sequences. One, (x_n), consisting of points in $[a,b]$, and the second, $\big(f(x_n)\big)$, consisting of points in $\mathbb{R}$.

First consider the second sequence, $\big(f(x_n)\big)$. This sequence was chosen so that $|f(x_n)| > n$ for each n. Therefore $\frac{1}{f(x_n)}$ is a null sequence by the Squeeze Rule (since $|\frac{1}{f(x_n)}| < \frac{1}{n}$ for each n and $\big(\frac{1}{n}\big)$ is a basic null sequence). Therefore $\big(f(x_n)\big)_{n=1}^{\infty}$ is divergent (by the remarks after Book D, Chapter 13, Worked Exercise 1.18). Moreover, we can apply the same reasoning to every subsequence $\big(f(x_{n_i})\big)$ of $\big(f(x_n)\big)$ (since $f(x_{n_i}) = f(x_j)$ for some $j \geq i$ and so $\big|\frac{1}{f(x_{n_i})}\big| \leq |\frac{1}{i}|$).

The Squeeze Rule is Lemma 1.15 of Chapter 13.

Now consider the first sequence, (x_n). Since this is a sequence of points in a closed interval, by Theorem 1.2 this sequence has a convergent subsequence, say (x_{n_i}), that is $d^{(1)}$-convergent to a point, say x, in the interval $[a,b]$. (We know that $x \in [a,b]$ by Book D, Chapter 16, Worked

Exercise 3.7.) Now f is continuous and so the limit of the sequence $(f(x_{n_i}))$ is $f(x)$, contradicting the fact that $(f(x_{n_i}))$ is not convergent. Therefore f is bounded and we have the result. ■

Theorem 1.4 *Extreme Value Theorem*

Let $f\colon [a,b] \to \mathbb{R}$ be a $(d^{(1)}_{[a,b]}, d^{(1)})$-continuous function on $[a,b]$. Then there are numbers $c, d \in [a,b]$ such that

$$f(c) \leq f(x) \leq f(d) \quad \text{for each } x \in [a,b].$$

Proof Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function on $[a,b]$. Then by Theorem 1.3 there is a real number M such that $|f(x)| \leq M$ for each $x \in [a,b]$. Therefore

$$-M \leq f(x) \leq M$$

for each $x \in [a,b]$.

We know that the image set $A = \{y \in \mathbb{R}\colon y = f(x) \text{ for some } x \in [a,b]\}$ is bounded above (by M) and so by the least upper bound property of the real numbers, the set A has a least upper bound, say L. Therefore we need to find a $d \in [a,b]$ with $f(d) = L$. We do this by constructing a sequence of points in $[a,b]$, (a_n), such that $f(a_n)$ gets closer and closer to L. We then use the fact that this sequence has a convergent subsequence to finish the proof.

See least upper bound property in the *Handbook*.

Since L is a least upper bound for the image set A, we have that for each n, $L - \frac{1}{n}$ is not an upper bound for A. Therefore we can find a point $a_n \in [a,b]$ with $L - \frac{1}{n} \leq f(a_n) \leq L$. This defines a sequence (a_n) in $[a,b]$. Therefore we now have two sequences: (a_n) in $[a,b]$, and $(f(a_n))$ in $A \subseteq \mathbb{R}$.

The sequence $(f(a_n))$ is $d^{(1)}$-convergent to L (when considered as a sequence in $\mathbb{R}$), but it may not be $d^{(1)}_A$-convergent (as L may not be in A).

For example, the sequence $\left(\frac{1}{n}\right)$, as a sequence in $\mathbb{R}$, is $d^{(1)}$-convergent to 0; but when considered as a sequence in $(0,1]$, $\left(\frac{1}{n}\right)$ is not $d^{(1)}_{(0,1]}$-convergent.

Now $[a,b]$ is a $d^{(1)}_{[a,b]}$-closed and bounded subset of $\mathbb{R}$, and so by Theorem 1.2, (a_n) has a convergent subsequence $\left(a_{n_i}\right)$. Set $d \in [a,b]$ equal to the limit of this convergent subsequence. (We know that $d \in [a,b]$ by Book D, Chapter 16, Worked Exercise 3.7.) Now since f is continuous, by Book D, Chapter 13, Definition 2.1, we know that the subsequence $f(a_{n_i}) \to f(d)$. Since every subsequence of a convergent sequence converges to the limit of the sequence (by Book D, Chapter 16, Lemma 1.13), we have $f(d) = L$ and f attains its upper bound.

Here we use the existence of a convergent subsequence for any sequence in $[a,b]$.

A similar argument shows that f attains its lower bound. ■

Remarks

In the proof of the Boundedness Theorem, we assumed that the function, f, was not bounded and used this assumption to construct a sequence in a closed and bounded interval. We then used the fact that this sequence had a convergent subsequence to get our contradiction.

In the proof of the Extreme Value Theorem, we used the Boundedness Theorem to find a bound on the function and then used the existence of this bound to construct a sequence in a closed and bounded interval. Again we used the fact that this sequence had a convergent subsequence to prove that the function attained its bounds. Therefore the important characteristic of the closed and bounded interval seems to be the fact that every sequence has a convergent subsequence.

2 Sequential compactness

In the proof of the Extreme Value Theorem, the key property of the $d^{(1)}$-closed and bounded interval $[a, b]$ that allows us to show that the function attains its bounds is the fact that the sequence (a_n) has a convergent subsequence. It is this property that we use to define *sequential compactness*, a generalisation of '$d^{(1)}$-closed and bounded'-ness in $\mathbb{R}$.

Definition 2.1 *d-sequentially compact*

A metric space (X, d) is **d-sequentially compact** if each sequence in X has a d-convergent subsequence (which is d-convergent to a point in X).

A subset $A \subseteq X$ is a **sequentially compact subset** if (A, d_A) is sequentially compact as a metric space.

Remarks

1. If the metric is clear from the context, then we suppress mention of the metric and say that X is *sequentially compact*.
2. If $A \subseteq X$ is a sequentially compact subset, then we sometimes refer to A as d_A-sequentially compact.
3. It is an essential part of this definition that the convergent subsequence converges to a point in the set under consideration.

Exercise 2.1

Prove that any sequentially compact subset of a metric space (X, d) is d-closed.

We wanted to generalise '$d^{(1)}$-closed and bounded'-ness in $\mathbb{R}$, and Exercise 2.1 shows that we have achieved part of this objective. The next result again links sequential compactness with '$d^{(1)}$-closed and bounded'-ness.

Corollary 2.2

Any $d^{(1)}$-closed and bounded subset of $\mathbb{R}$ is sequentially compact.

Proof This is an immediate corollary of Theorem 1.2. ■

In fact, it turns out that the sequentially compact subspaces of $\mathbb{R}$ are precisely those which are $d^{(1)}$-closed and bounded.

Theorem 2.3 *The Heine–Borel Theorem for* $\mathbb{R}$

A subset of $\mathbb{R}$ is $d^{(1)}$-sequentially compact if, and only if, it is $d^{(1)}$-closed and bounded.

Proof Corollary 2.2 states that any $d^{(1)}$-closed and bounded subset of $\mathbb{R}$ is sequentially compact. We need to prove that any sequentially compact subset of $\mathbb{R}$ (with the Euclidean metric) is both $d^{(1)}$-closed and bounded.

By Exercise 2.1, a $d^{(1)}$-sequentially compact subset of $\mathbb{R}$ is $d^{(1)}$-closed. The proof that A is bounded is by contradiction. Assuming that A is not bounded, it is possible to construct a sequence in A which does not have a convergent subsequence. We leave this for you as Exercise 2.2. Therefore, by contradiction, A is bounded and we have the result. ■

Exercise 2.2

Show that every sequentially compact subset, A, of $\mathbb{R}$ (with the usual Euclidean metric), is bounded by assuming that A is not bounded and constructing a sequence in A that does not have a convergent subsequence.

2.1 Examples of sequentially compact spaces

We know, from Theorem 2.3 that the sequentially compact subsets of $\mathbb{R}$ are precisely the $d^{(1)}$-closed and bounded subsets of $\mathbb{R}$. We now look at some other examples and non-examples.

Exercise 2.3

Is K_C, the middle-third Cantor set (with the metric induced by the Euclidean metric on $\mathbb{R}$), sequentially compact? Prove your answer.

The middle-third Cantor set is defined in Book D, Chapter 16, Example 1.19.

Remember that to check that a set is sequentially compact you must show that *every* sequence has a convergent subsequence.

Exercise 2.4

Let (X, d) be a metric space where X is a finite set and d is any metric. Show that X is sequentially compact.

To show that a set is not sequentially compact just requires finding *one* sequence without a convergent subsequence.

Worked Exercise 2.4

Let

$$A = \{f \in C[0,1] : |f(x)| \leq 1 \text{ for all } x \in [0,1]\}.$$

Show that A is not sequentially compact for $d_{\max}$.

See Book D, Chapter 15, Definition 5.1.

Solution

We need to find just one sequence of functions in the set A that has no subsequence converging to a point in the set A.

With this in mind, we consider the sequence (f_n) in A given by

$$f_n(x) = \begin{cases} 1 - nx, & \text{if } 0 \leq x < \frac{1}{n}, \\ 0, & \text{if } \frac{1}{n} \leq x \leq 1. \end{cases}$$

The first few functions in the sequence are shown in Figure 2.1.

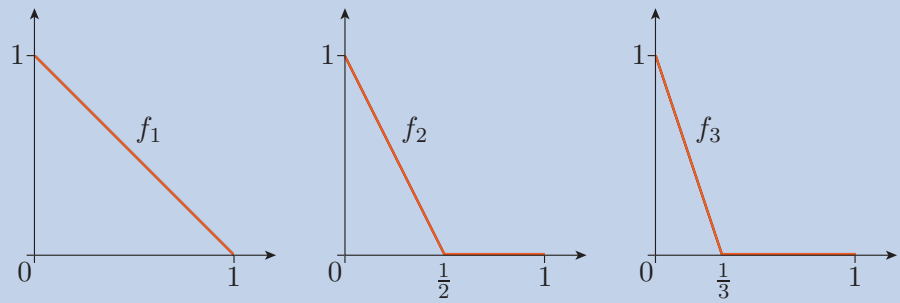


Figure 2.1

We know that, for $x \in [0,1]$, the limit as n tends to infinity of $f_n(x)$ is given by:

$$f_n(x) \to \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{if } 0 < x \leq 1. \end{cases}$$

Thus the pointwise limit of the sequence is a discontinuous function. Moreover, any subsequence of (f_n) has the same pointwise limit. Hence no subsequence can converge to a continuous function and A is not sequentially compact.

We investigated this sequence of functions and calculated the limit in Book D, Chapter 15, Worked Exercise 5.9.

Exercise 2.5

Let

$$A = \{f \in C[0,1] : |f(x) - f(y)| \leq |x - y| \text{ for all } x, y \in [0,1]\}.$$

Show that A is not sequentially compact for $d_{\max}$.

Hint: find a sequence (f_n) of functions in A such that (f_n) and any subsequence of (f_n) have no pointwise limit, and then make use of results from Book D, Chapter 15.

2.2 Subsets, images and products of sequentially compact spaces

Suppose that (X, d) is a d-sequentially compact metric space. It is natural to ask whether there are any subsets A of X which are necessarily d_A-sequentially compact. We know from Exercise 2.4 that any finite subset is sequentially compact. The next result tells us that any d-closed subset A of X is d_A-sequentially compact.

Theorem 2.5

Let $A \subseteq X$ where (X, d) is a d-sequentially compact metric space. Then A is d-closed if, and only if, A is d_A-sequentially compact.

Proof We first show that a d-closed subset is sequentially compact.

Let A be a d-closed subset of a sequentially compact metric space (X, d) and let (a_n) be a sequence in A. We need to find a subsequence of (a_n) which is d_A-convergent. Since X is sequentially compact, and (a_n) is a sequence in X, we know that (a_n) has a d-convergent subsequence, say (a_{n_i}); we need to prove that the limit of this subsequence is in A. However, since A is closed, by definition, if we have any d-convergent sequence with terms in A, then the limit must lie in A. The subsequence (a_{n_i}) is such a sequence and so the limit of this must lie in A. Therefore (a_{n_i}) is d_A-convergent and A is sequentially compact.

By Exercise 2.1, any sequentially compact subset of any metric space (X, d) is d-closed, and so sequentially compact subsets of sequentially compact metric spaces are certainly closed. ■

Remark

In particular, by taking $A = X$, we see that sequentially compact sets are closed.

Since convergence is preserved by continuous functions, it is no surprise that the continuous image of a sequentially compact set is sequentially compact.

Lemma 2.6

Let (X, d) and (Y, e) be metric spaces with X being sequentially compact. Let $f: X \to Y$ be a (d, e)-continuous function. Then the image, $f(X)$, is sequentially compact.

Proof This is Exercise 2.6. ■

Exercise 2.6

Prove Lemma 2.6.

This result tells us that sequential compactness is a topological invariant.

Corollary 2.7

Sequential compactness is preserved by homeomorphism.

Proof Let (X, d) and (Y, e) be homeomorphic metric spaces and let $f: X \to Y$ be a homeomorphism between them. The map $f: X \to Y$ is an onto and (d, e)-continuous map. Therefore if X is sequentially compact then Lemma 2.6 implies that Y must be sequentially compact. Also $f^{-1}: Y \to X$ is an onto (e, d)-continuous map from Y to X so, by the same lemma, X is sequentially compact if Y is sequentially compact. Therefore X is sequentially compact if, and only if, Y is sequentially compact. ■

We know that many properties are preserved by taking products of metric spaces. Sequential compactness is one of these properties.

Theorem 2.8 *Tikhonov's Theorem*

Let (X, d) and (Y, d') be two sequentially compact metric spaces. Then $(X \times Y, e)$ is sequentially compact (where e is a product metric).

Proof We need to show that every sequence in $X \times Y$ contains an e-convergent subsequence. Let (x_n, y_n) be a sequence in $X \times Y$. First consider (x_n) as a sequence in X. Since X is sequentially compact, this sequence has a convergent subsequence, which we write as (x_{n_i}). Now consider the sequence (y_{n_i}) in Y. This is not, of course, necessarily convergent but, since Y is sequentially compact, it also has a convergent subsequence, which we write as $(y_{n_{i_j}})$.

Now consider the sequence $(x_{n_{i_j}}, y_{n_{i_j}})$. By Book D, Chapter 15, Theorem 3.15, this sequence is e-convergent to a point (x, y) in $X \times Y$ if the component sequence $\left(x_{n_{i_j}}\right)$ is d-convergent to x in X and the component sequence $\left(y_{n_{i_j}}\right)$ is d'-convergent to y in Y. We know that

$(y_{n_{i_j}})$ is convergent to a point in Y and $(x_{n_{i_j}})$ is a subsequence of the convergent sequence (x_{n_i}), and so converges to the same point in X (by Book D, Chapter 16, Lemma 1.13) – see Figure 2.2.

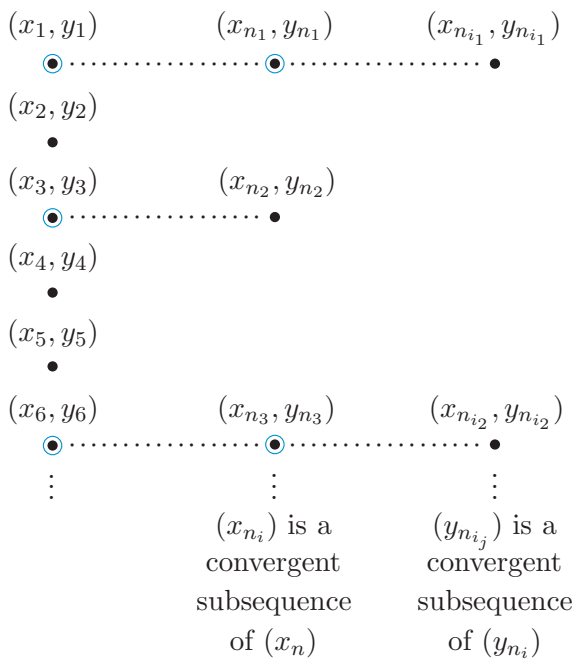


Figure 2.2

Therefore $(x_{n_{i_j}}, y_{n_{i_j}})$ is a convergent subsequence of our arbitrary sequence (x_n, y_n) and $X \times Y$ is sequentially compact. ■

2.3 Sequentially compact subsets of $\mathbb{R}^n$

We know from Corollary 2.2 that any $d^{(1)}$-closed and bounded subset of $\mathbb{R}$ (with the usual Euclidean metric) is sequentially compact. We can use Theorem 2.8 to show that any $d^{(n)}$-closed and bounded subset of $\mathbb{R}^n$ is sequentially compact, but first we need to say what we mean by the term *bounded* when we are in a metric space other than $(\mathbb{R}, d^{(1)})$.

Boundedness is defined for subsets of $\mathbb{R}$ (with the Euclidean metric) in Book D, Chapter 13, Definition 2.14: a set $A \subset \mathbb{R}$ is *bounded* if there exists a real number M such that $|x| \leq M$ for each $x \in A$.

Definition 2.9 *Bounded set*

Let (X, d) be a metric space and let $A \subseteq X$. Then A is **bounded** if there is $M > 0$ such that $d(x, y) \leq M$ for all $x, y \in A$. If we want to emphasise the metric, we say that A is **d-bounded**.

Intuitively, a subset A of a metric space is bounded if we can find a number M such that all the points in A are within M of one another. This number M can be as large as we like, but it must be finite.

Worked Exercise 2.10

Show that the definition of a bounded subset of $\mathbb{R}$ given in Book D, Chapter 13, Definition 2.14 is equivalent to Definition 2.9 for the Euclidean metric on $\mathbb{R}$.

Solution

Let $A \subseteq \mathbb{R}$ be bounded. If $A = \emptyset$, then there is nothing to prove, so suppose $A \neq \emptyset$. Suppose that A is bounded as defined in Definition 2.9. Then there exists $M > 0$ such that $d^{(1)}(x, y) \leq M$ for all $x, y \in A$. Hence, fixing some $y \in A$ and using the Triangle Inequality, we find for any $x \in A$,

$$|x| = |x - y + y| \leq |x - y| + |y| \leq M + |y|.$$

Thus A is bounded according to Definition 2.14 of Chapter 13.

Now suppose that there exists $M > 0$ such that $|x| \leq M$ for all $x \in A$. Then, by the Triangle Inequality,

$$|x - y| \leq |x| + |y| \leq M + M = 2M,$$

for all $x, y \in A$. Thus A is bounded as defined in Definition 2.9.

Exercise 2.7

The projection functions are defined in Book D, Chapter 14, Definition 1.18.

Show that if $A \subseteq \mathbb{R}^n$ is $d^{(n)}$-bounded, and p_i is the ith projection function then the image of A under this function, $p_i(A) \in \mathbb{R}$ is $d^{(1)}$-bounded.

The definition of bounded for a subset A of a metric space (X, d) is equivalent to saying that we can enclose A within a ball (as we ask you to show in the next exercise).

Exercise 2.8

Show that the definition of bounded as given in Definition 2.9 is equivalent to the statement that A can be enclosed within an open ball $B_d(a, r)$ for some $a \in A$ and $r > 0$.

For $\mathbb{R}^n$ we can show that any closed and bounded set is $d^{(n)}$-sequentially compact.

Proposition 2.11

Any $d^{(n)}$-closed and bounded subset of $\mathbb{R}^n$ is $d^{(n)}$-sequentially compact.

Proof Suppose that $K \subseteq \mathbb{R}^n$ is $d^{(n)}$-closed and bounded. Since K is bounded, by Exercise 2.7 $p_j(K)$ is bounded for each $j = 1, 2, \ldots, n$, where p_j is the usual projection function onto the jth coordinate axis. So we can find $M_j > 0$ such that

$$p_j(K) \subseteq [-M_j, M_j], \quad j = 1, 2, \ldots, n.$$

Hence

$$K \subseteq [-M_1, M_1] \times [-M_2, M_2] \times \cdots \times [-M_n, M_n]$$

(Figure 2.3). Each of the intervals $[-M_j, M_j]$ is a sequentially compact subset of $\mathbb{R}$. Thus, by Tikhonov's Theorem, the product of the intervals is also sequentially compact. Hence K is a closed subset of a sequentially compact set, and so is sequentially compact, by Theorem 2.5. ■

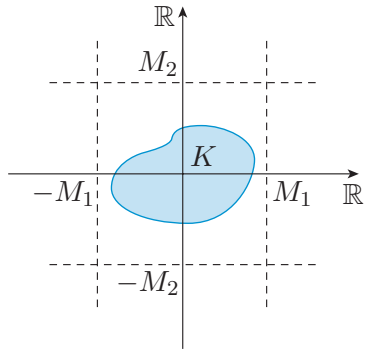


Figure 2.3

Exercise 2.9

Show that S^3, the unit sphere in $\mathbb{R}^4$, given by $\{(w, x, y, z) \in \mathbb{R}^4 : w^2 + x^2 + y^2 + z^2 = 1\}$, is sequentially compact (for the Euclidean metric). (*Hint*: you may find Book D, Chapter 16, Theorem 5.2 helpful.)

We now know that all $d^{(n)}$-closed and bounded subsets of $\mathbb{R}^n$ are sequentially compact. We next show that these are the only sequentially compact subsets of $\mathbb{R}^n$.

Theorem 2.12 *The Heine–Borel Theorem*

A subset of $\mathbb{R}^n$ is $d^{(n)}$-sequentially compact if, and only if, it is $d^{(n)}$-closed and bounded.

Proof We know from Proposition 2.11 that any $d^{(n)}$-closed and bounded subset of $\mathbb{R}^n$ is sequentially compact. Therefore it is sufficient to prove that any $d^{(n)}$-sequentially compact subset of $\mathbb{R}^n$ is $d^{(n)}$-closed and bounded.

We know from Exercise 2.1 that a sequentially compact subset of any metric space (X, d) is d-closed, so this is certainly true for a subset of $\mathbb{R}^n$. Therefore it is sufficient to prove that A is bounded.

The proof that A is bounded is by contradiction. Assume that A is not bounded. Then A cannot be empty and so we may assume that there is an element $a_1 \in A$. By Exercise 2.8 we cannot find an $M > 0$ such that $B_{d^{(n)}}(a_1, M) \supseteq A$. Therefore for each $n \in \mathbb{N}$ we can find an $a_n \in A$ with $d(a_1, a_n) > n$. This gives us a sequence (a_n) in A.

Since $a_n \in A$, and A is sequentially compact, (a_n) has a convergent subsequence (a_{n_i}). Suppose that a is the limit of (a_{n_i}). Then we can find

$I \in \mathbb{N}$ so that for $i > I$, $d(a_{n_i}, a) < 1$. Thus for $i > I$, on using the Reverse Triangle Inequality for d, we find that

$$\begin{aligned} d(a_1, a) &\geq |d(a_1, a_{n_i}) - d(a_{n_i}, a)| \\ &\geq d(a_1, a_{n_i}) - 1 \\ &\geq n_i - 1, \text{ for each } i > I. \end{aligned}$$

But $n_i \to \infty$ and so $d(a_1, a)$ is not finite, which is impossible; we conclude that A is bounded. ■

As promised we now generalise the Extreme Value Theorem.

2.4 The Generalised Extreme Value Theorem

We can now prove a general version of the Extreme Value Theorem.

Theorem 2.13 *General Extreme Value Theorem*

Let (X, d) be a non-empty sequentially compact metric space, and let $f\colon X \to \mathbb{R}$ be $(d, d^{(1)})$-continuous. Then there are $c, d \in X$ such that

$$f(c) \leq f(x) \leq f(d) \quad \text{for all } x \in X.$$

Proof Since X is sequentially compact and f is continuous, Lemma 2.6 implies that $f(X)$ is a sequentially compact subset of $\mathbb{R}$. Hence, by Theorem 2.3, $f(X)$ is a closed and bounded subset of $\mathbb{R}$.

See upper bound and lower bound in the *Handbook*.

Since the image $f(X)$ is bounded, it has an upper bound and a lower bound. Therefore, by the least upper bound property of the real numbers, the set $f(X)$ has a least upper bound M and a greatest lower bound m. For any $r > 0$, since $M - \frac{r}{2}$ is not an upper bound for $f(X)$, we can find a point $p \in f(X)$ with $M - \frac{r}{2} \leq f(p) \leq M$. Therefore we can deduce that $B_{d^{(1)}}(M, r)$ must intersect $f(X)$ for any $r > 0$ and so, by Book D, Chapter 16, Worked Exercise 3.7, we see that M is a closure point of $f(X)$. But $f(X)$ is closed, and so we deduce that $M \in f(X)$. A similar argument shows that $m \in f(X)$. Thus there are $c \in X$ and $d \in X$ with $f(c) = m$ and $f(d) = M$. It follows from the definition of m and M that $f(c) \leq f(x) \leq f(d)$ for all $x \in X$. ■

It follows from Theorem 2.13 that any $(d^{(n)}, d^{(1)})$-continuous function from a sequentially compact subset of $\mathbb{R}^n$ to $\mathbb{R}$ has a least upper bound and a greatest lower bound, and attains them both; in other words, it has a maximum value and a minimum value.

If the function is $(d^{(n)}, d^{(1)})$-continuous but the domain is not sequentially compact, then this need not be the case.

Worked Exercise 2.14

Let $f: \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x, y) = \sqrt{x^2 + y^2}.$$

Show that f attains a maximum value on $B_{d^{(2)}}[(0, 0), 1]$; see Figure 2.4.

Solution

In order to use Theorem 2.13 to deduce the existence of a maximum value, we must verify that f is $(d^{(2)}, d^{(1)})$-continuous on its domain and that $B_{d^{(2)}}[(0, 0), 1]$ is a sequentially compact set.

We saw in Book D, Chapter 13, Worked Exercise 3.17 that f is continuous on the whole of $\mathbb{R}^2$, and so, by the Restriction Rule, is continuous on $B_{d^{(2)}}[(0, 0), 1]$.

The ball $B_{d^{(2)}}[(0, 0), 1]$ is closed and bounded. Hence, by Theorem 2.12, it is sequentially compact.

Thus both conditions of the General Extreme Value Theorem are satisfied, and we conclude that the function f attains a maximum value on $B_{d^{(2)}}[(0, 0), 1]$.

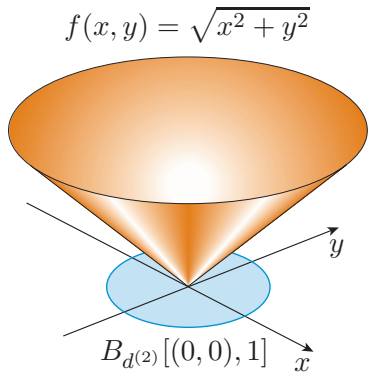


Figure 2.4

Exercise 2.10

Let $f: \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x, y) = y \sin(\pi x) + x \cos(\pi y).$$

Show that f attains a minimum value on the set $\{(x, y) : 0 \leq x \leq 1, \text{ and } 0 \leq y \leq 1\}$.

The General Extreme Value Theorem is a powerful result since it guarantees the existence of a maximum value and a minimum value of a continuous function on a sequentially compact domain without our having to know where these values are.

We now turn our attention to subsets of the function space $C[0, 1]$.

3 Functions and sequential compactness

In this section we investigate the relationship between sequential compactness and functions in $(C[0, 1], d_{\max})$, with the goal of finding sufficient conditions to ensure that a given subset of $C[0, 1]$ is sequentially compact.

We begin by examining in a little more detail what it means for a function from $[0,1]$ to $\mathbb{R}$ to be continuous. We discover that the fact that $[0,1]$ is sequentially compact means that f satisfies a stronger notion of continuity, known as *uniform continuity*. This will lead us to identifying one of the conditions needed to guarantee sequential compactness of a subset of $C[0,1]$.

3.1 Sequentially compact subsets of $\mathbb{R}$ and continuous functions

We start by finding an alternative formulation of what it means for a function from a subset of $\mathbb{R}$ to $\mathbb{R}$ to be continuous (for the Euclidean metrics).

Proposition 3.1

Let $A \subseteq \mathbb{R}$ and $f\colon A \to \mathbb{R}$. Then f is (d_A, d)-continuous at $a \in A$ if, and only if, for each $\varepsilon > 0$, there is $\delta > 0$ such that

$$|f(x) - f(a)| < \varepsilon, \text{ whenever } x \in A \text{ and } |x - a| < \delta.$$

(Here d denotes the Euclidean metric for $\mathbb{R}$ and d_A denotes the corresponding induced metric for A.)

Proof *Only if*

Let $a \in A$ and suppose that f is (d_A, d)-continuous at $a \in A$.

Let $\varepsilon > 0$ be given. We suppose, for a contradiction, that for each $\delta > 0$, we can find $x \in B_{d_A}(a, \delta)$ for which $|f(x) - f(a)| \geq \varepsilon$. Then, in particular, for each $n \in \mathbb{N}$, we can find $x_n \in B_{d_A}(a, \frac{1}{n})$ for which $|f(x_n) - f(a)| \geq \varepsilon$. Thus for this sequence (x_n), we have that it converges to a and yet $|f(x_n) - f(a)| \geq \varepsilon > 0$ for each $n \in \mathbb{N}$, meaning that the sequence $(f(x_n))$ does not converge to $f(a)$ – this contradicts the assumption that f is continuous at a.

If

For the opposite implication, let $a \in A$ and suppose that for each $\varepsilon > 0$, there is $\delta > 0$ such that $|f(x) - f(a)| < \varepsilon$, whenever $x \in A$ and $|x - a| < \delta$.

Let (x_n) be a sequence in A that d_A-converges to a. Our aim is to show that $f(x_n)$ converges to $f(a)$ as $n \to \infty$.

Let $\varepsilon > 0$ be given. Then, by our assumption, we can find $\delta > 0$ for which $|f(x) - f(a)| < \varepsilon$, whenever $d_A(x, a) < \delta$. However, since $d_A(x_n, a) \to 0$ as $n \to \infty$, we can find $N \in \mathbb{N}$ so that $d_A(x_n, a) < \delta$ for $n > N$, and so for $n > N$, $d(f(x_n), f(a)) < \varepsilon$. In other words, for each $\varepsilon > 0$, there is $N \in \mathbb{N}$ so that $d(f(x_n), f(a)) < \varepsilon$ for $n > N$, which means that $f(x_n)$ converges to $f(a)$ as $n \to \infty$, and we conclude that f is continuous at a, as required. ■

Remark

This equivalent formulation of continuity is known as the ε-δ **definition of continuity** – you will likely have met it in your earlier studies. It is discussed in M208.

An examination of the proof reveals that there is no particular requirement to restrict our attention to a subset of the real line, and a similar result can be proved in general metric spaces.

Theorem 3.2

Let (X, d) and (Y, e) be metric spaces and let $f\colon X \to Y$. Then f is (d, e)-continuous at $a \in X$ if, and only if, for each $\varepsilon > 0$, there is $\delta > 0$ so that $e(f(x), f(a)) < \varepsilon$, whenever $d(x, a) < \delta$.

Proof See Exercise 3.1. ■

Exercise 3.1

Prove Theorem 3.2.

These results are only about continuity and we have not made use of sequential compactness – in our next results, we look at what happens in the case that the domain of the function is sequentially compact (and so, in particular, there is a metric available). Our ultimate goal in this section is to understand what it means for a subset of $C[0, 1]$ to be sequentially compact. We know from Corollary 2.2 that the interval $[0, 1]$ (and any other closed bounded interval $[a, b]$) is itself a sequentially compact subset of $\mathbb{R}$ for the Euclidean metric, and so we start by looking at a single continuous function with this domain.

Theorem 3.3

Let $f\colon [0, 1] \to \mathbb{R}$ be a $(d^{(1)}_{[0,1]}, d^{(1)})$-continuous function. Then for each $\varepsilon > 0$, there is $\delta > 0$ such that for each $x, y \in [0, 1]$, $|f(x) - f(y)| < \varepsilon$, whenever $|x - y| < \delta$.

Remarks

1. You should contrast this with the statement of Proposition 3.1 to see how these results differ: in Proposition 3.1, the choice of δ will in general depend on both ε and the point a, whereas here the same δ will do for *all* points x in $[0, 1]$.
2. A similar result holds if the interval $[0, 1]$ is replaced by a closed bounded interval $[a, b]$.

Proof We prove the theorem using proof by contradiction. Suppose that there is $\varepsilon > 0$ such that for each $\delta > 0$ there are $x, y \in [0,1]$ for which $|x - y| < \delta$ and yet $|f(x) - f(y)| \geq \varepsilon$.

Then, in particular, for each $n \in \mathbb{N}$, we can find $x_n, y_n \in [0,1]$ for which $|x_n - y_n| < \frac{1}{n}$ and yet $|f(x_n) - f(y_n)| \geq \varepsilon$.

Consider the sequence (x_n). Since each term lies in the sequentially compact set $[0,1]$, there is a convergent subsequence $(x_{n_i})_{i \in \mathbb{N}}$, with limit x, say, that lies in $[0,1]$. Since f is continuous, we know that $f(x_{n_i}) \to f(x)$ as $n \to \infty$.

Let us look at the corresponding subsequence (y_{n_i}) of (y_n). Then we have $|x_{n_i} - y_{n_i}| < \frac{1}{n_i} \leq \frac{1}{i}$ for each $i \in \mathbb{N}$, and so the Triangle Inequality for modulus implies that

$$|x - y_{n_i}| \leq |x - x_{n_i}| + |x_{n_i} - y_{n_i}| \leq |x - x_{n_i}| + \frac{1}{i}.$$

But $|x - x_{n_i}| \to 0$ as $i \to \infty$ and so, by the Squeeze Rule, we deduce that $y_{n_i} \to x$ as $i \to \infty$. Hence, since f is continuous, $f(y_{n_i}) \to f(x)$ as $i \to \infty$.

However, for each $i \in \mathbb{N}$, the definition of x_{n_i} and y_{n_i}, together with the Triangle Inequality for modulus, implies that

$$\begin{aligned} 0 < \varepsilon &\leq |f(x_{n_i}) - f(y_{n_i})| \\ &\leq |f(x_{n_i}) - f(x)| + |f(x) - f(y_{n_i})| \to 0 \text{ as } i \to \infty, \end{aligned}$$

which is a contradiction. We conclude that the theorem holds. ■

If you look closely at the proof of this result, you will see that the only property about $[0,1]$ that we made use of is that it is sequentially compact: an almost identical argument shows that a similar result holds in the case when the domain of the continuous function is sequentially compact.

Theorem 3.4

Let (X, d) and (Y, e) be metric spaces and suppose $f : X \to Y$ is (d, e)-continuous on X.

If X is d-sequentially compact, then for each $\varepsilon > 0$, there is $\delta > 0$ such that for each $x, y \in X$, $e(f(x), f(y)) < \varepsilon$ whenever $d(x, y) < \delta$.

Proof See Exercise 3.2. ■

Exercise 3.2

Prove Theorem 3.4.

This property (of continuous functions on sequentially compact sets) is important enough to have been given a name: *uniform continuity*.

Definition 3.5 *Uniform continuity*

Let (X, d) and (Y, e) be metric spaces and suppose that $f\colon X \to Y$ satisfies the following condition: for each $\varepsilon > 0$, there is $\delta > 0$ such that for each $x, y \in X$, if $d(x, y) < \delta$, then $e(f(x), f(y)) < \varepsilon$.

Then f is (d, e)-**uniformly continuous**.

Remarks

1. In particular, Theorem 3.3 tells us that every function in $C[0, 1]$ is uniformly continuous on $[0, 1]$ for the Euclidean metrics.
2. Notice that this definition does *not* require that the underlying metric space is sequentially compact.
3. The fact that a real-valued continuous function defined on a bounded closed interval (with the Euclidean metrics) is uniformly continuous can be used to prove that it is (Riemann) integrable.

This is proved in unit AB3 of M208.

We finish this subsection by recording a useful lemma that shows that uniformly continuous functions are always continuous.

Lemma 3.6

Let (X, d) and (Y, e) be metric spaces and suppose that $f\colon X \to Y$. If f is (d, e)-uniformly continuous on X, then f is (d, e)-continuous on X.

Exercise 3.3

Prove Lemma 3.6.

Exercise 3.4

Give an example of a continuous function $f\colon \mathbb{R} \to \mathbb{R}$ that is not uniformly continuous (for the Euclidean metrics on both its domain and codomain).

3.2 Sequentially compact subsets of $C[0, 1]$

In Worked Exercise 2.4 and Exercise 2.5, we investigated sets of functions in $C[0, 1]$ that turned out not to be sequentially compact. In this subsection we state, and show how to use, a result that gives us sufficient conditions to guarantee that a subset of $C[0, 1]$ *is* sequentially compact.

Before we state our main result for this section, we need a couple of definitions, both of which apply to a *set* of functions.

Our first definition gives a notion of boundedness for collections of functions.

Definition 3.7 *Pointwise bounded*

Let A be a set of functions from X to $\mathbb{R}$. We say that the set A is **pointwise bounded** if for each $x \in X$, there is $M_x \geq 0$ such that for each $f \in A$, $|f(x)| \leq M_x$.

So each $f \in A$ has domain X and codomain $\mathbb{R}$.

Remarks

1. The constant M_x may depend on the point $x \in X$ chosen but cannot depend on the choice of function from A.
2. This definition may remind you a little of the definition of pointwise convergence that we gave in Book D, Chapter 15 inasmuch as it does not require there to be a metric defined for the domain of the functions and so, in particular, there are no assumptions about whether the functions are continuous.
3. If the set A consists of just a single function, $A = \{f\}$ say, then this definition is automatically satisfied – a set consisting of one function is *always* pointwise bounded. In particular, saying the set $A = \{f\}$ is pointwise bounded is *not* the same as saying that the function f is bounded.

Exercise 3.5

Let $g\colon \mathbb{R} \to \mathbb{R}$ be a function and let A_g be the set consisting of all functions $f\colon \mathbb{R} \to \mathbb{R}$ such that $|f(x)| \leq |g(x)|$ for each $x \in \mathbb{R}$. Show that A is pointwise bounded.

Exercise 3.6

Let $A = \{f\colon [0,1] \to \mathbb{R} : f \text{ is constant}\}$.

Show that A is not pointwise bounded.

Exercise 3.7

Let (X, d) be a metric space, let $L \geq 0$, and fix $a \in X$ and $b \in \mathbb{R}$. Define

$$A = \{f\colon X \to \mathbb{R} : f(a) = b \text{ and } f \text{ is } (d, d^{(1)})\text{-Lipschitz with Lipschitz constant } L\}.$$

Lipschitz functions are defined in Book D, Chapter 15. A function $f\colon X \to \mathbb{R}$ is $(d, d^{(1)})$-Lipschitz with Lipschitz constant L if $|f(x) - f(y)| \leq Ld(x, y)$ for each $x, y \in X$.

Show that A is a pointwise bounded set.

Exercise 3.8

Let (X, d) be a metric space and let $L \geq 0$. Define

$$A = \{f\colon X \to \mathbb{R} : f \text{ is } (d, d^{(1)})\text{-Lipschitz with Lipschitz constant } L\}.$$

Determine whether A is a pointwise bounded set.

Exercise 3.9

Show that if A is a sequentially compact subset of $(C[0,1], d_{\max})$, then it is pointwise bounded.

Our second definition extends the notion of uniform continuity to a collection of functions.

Definition 3.8 *Equicontinuity*

Let (X, d) and (Y, e) be metric spaces and let A be a set of functions from X to Y.

The set A is a (d, e)**-equicontinuous set of functions** if for each $\varepsilon > 0$, there is a $\delta > 0$ such that for each $f \in A$ and every $x, y \in X$, $e(f(x), f(y)) < \varepsilon$ whenever $d(x, y) < \delta$.

Remarks

1. Although we give the general definition here, we shall only use it for collections of continuous functions from $[0,1]$ to $\mathbb{R}$.
2. We usually suppress mention of the metrics when it is clear from the context and just say that A is an equicontinuous set (or family) of functions.
3. If you compare this definition with that of the definition of uniform continuity (for a single function), Definition 3.5, you will see that the main difference is that the same δ has to work for all functions f in the set A. In particular, if A consists of just a single function $A = \{f\}$, then saying that A is equicontinuous is the same as saying that f is uniformly continuous.

Worked Exercise 3.9

Let $L \geq 0$ be given and let

$$A = \{f \colon [0,1] \to \mathbb{R} : f \text{ is } (d^{(1)}_{[0,1]}, d^{(1)})\text{-Lipschitz with Lipschitz constant } L\}.$$

Show that A is an equicontinuous set of functions (for the Euclidean metrics on $[0,1]$ and $\mathbb{R}$).

Solution

Since each $f \in A$ is Lipschitz with Lipschitz constant L, we have for each $x, y \in [0,1]$,

$$|f(x) - f(y)| \leq L|x - y|.$$

The '$L+1$' is to cover the case $L=0$.

Hence if $|x-y|<\delta$, then $|f(x)-f(y)|\leq L\delta$. Thus if $\varepsilon>0$ is given, then on choosing $\delta=\varepsilon/(L+1)$, we find that if $|x-y|<\delta$, then for any $f\in A$, $|f(x)-f(y)|<\varepsilon$. Hence A is an equicontinuous set of functions.

Exercise 3.10

Let (X,d) and (Y,e) be metric spaces and let A be a (d,e)-equicontinuous set of functions from X to Y. Show that if $B\subseteq A$, then B is also a (d,e)-equicontinuous set of functions.

Our next lemma shows that functions in an equicontinuous set are continuous.

Lemma 3.10

Let (X,d) and (Y,e) be metric spaces and let A be a (d,e)-equicontinuous set of functions from X to Y. Then each $f\in A$ is (d,e)-uniformly continuous and hence (d,e)-continuous on X.

Exercise 3.11

Prove Lemma 3.10.

We now identify a class of subsets of $C[0,1]$ that are equicontinuous (for the Euclidean metrics on both domain and codomain).

Proposition 3.11

If A is a $d_{\max}$-sequentially compact subset of $C[0,1]$, then A is equicontinuous.

Proof We give a proof by contradiction: suppose that A is a $d_{\max}$-sequentially compact subset of $C[0,1]$ that is not equicontinuous.

Then there is $\varepsilon>0$ such that for each $\delta>0$, there is $f\in A$ and $x,y\in[0,1]$ such that $|f(x)-f(y)|\geq\varepsilon$ and $|x-y|<\delta$.

In particular, taking $\delta=\frac{1}{n}$, we can find $f_n\in A$ and $x_n,y_n\in[0,1]$ such that $|f_n(x_n)-f_n(y_n)|\geq\varepsilon$ and $|x_n-y_n|\leq\frac{1}{n}$. Thus we have a sequence of functions (f_n) and two sequences of points (x_n) and (y_n). Since A is sequentially compact, (f_n) has a $d_{\max}$-convergent subsequence, (f_{n_k}), say, converging to a function $f\in A$. In particular, f is continuous.

Consider now the corresponding sequence of points (x_{n_k}), which lies within the interval $[0,1]$. Since, by Theorem 1.2, $[0,1]$ is $d^{(1)}$-sequentially compact, we can find a subsequence of (x_{n_k}), $(x_{n_{k_i}})_{i=1}^{\infty}$, say, that converges to $x \in [0,1]$. Moreover, since $|x_n - y_n| < \frac{1}{n}$ for each $n \in \mathbb{N}$, we conclude that the sequence $(y_{n_{k_i}})_{i=1}^{\infty}$ also converges to x. Now choose $I \in \mathbb{N}$ so that $d_{\max}(f_{n_k}, f) < \frac{\varepsilon}{4}$ for $k > I$. Then since $k_i \geq i$, we have that for $i > I$,

$$\begin{aligned}
\varepsilon &\leq |f_{n_{k_i}}(x_{n_{k_i}}) - f_{n_{k_i}}(y_{n_{k_i}})| \\
&= |f_{n_{k_i}}(x_{n_{k_i}}) - f(x_{n_{k_i}}) + f(x_{n_{k_i}}) - f(x) + f(x) - f(y_{n_{k_i}}) + f(y_{n_{k_i}}) - f_{n_{k_i}}(y_{n_{k_i}})| \\
&\leq |f_{n_{k_i}}(x_{n_{k_i}}) - f(x_{n_{k_i}})| + |f(x_{n_{k_i}}) - f(x)| + |f(x) - f(y_{n_{k_i}})| \\
&\qquad + |f(y_{n_{k_i}}) - f_{n_{k_i}}(y_{n_{k_i}})| \\
&\leq d_{\max}(f, f_{n_{k_i}}) + |f(x_{n_{k_i}}) - f(x)| + |f(x) - f(y_{n_{k_i}})| + d_{\max}(f, f_{n_{k_i}}) \\
&< \tfrac{\varepsilon}{2} + |f(x_{n_{k_i}}) - f(x)| + |f(x) - f(y_{n_{k_i}})|.
\end{aligned}$$

Since f is continuous and $x_{n_{k_i}} \to x$, we can choose $J > I$ so that for $i > J$, $|f(x_{n_{k_i}}) - f(x)| < \frac{\varepsilon}{4}$. Hence for $i > J$, we have

$$\varepsilon < \frac{3\varepsilon}{4} + |f(x) - f(y_{n_{k_i}})|,$$

which rearranges to give $|f(x) - f(y_{n_{k_i}})| > \frac{\varepsilon}{4}$ for $i > J$. But $y_{n_{k_i}} \to x$ as $i \to \infty$, and so we conclude that f is not continuous at x, which is a contradiction. ■

Summarising our results for sequentially compact subsets of $C[0,1]$ (with the $d_{\max}$ metric), we have:

1. sequentially compact subsets of $C[0,1]$ are $d_{\max}$-closed (by Theorem 2.5)
2. sequentially compact subsets of $C[0,1]$ are pointwise bounded (by Exercise 3.9)
3. sequentially compact subsets of $C[0,1]$ are equicontinuous (by Proposition 3.11).

The following theorem, which is the main result of this section, shows that the converse also holds: these three properties are sufficient to identify sequentially compact subsets of $(C[0,1], d_{\max})$.

Theorem 3.12 *The Arzelà–Ascoli Theorem*

Let A be a subset of $C[0,1]$. Suppose that:

1. A is $d_{\max}$-closed
2. A is a pointwise bounded set of functions
3. A is an equicontinuous set of functions.

Then A is $d_{\max}$-sequentially compact.

This theorem is named after the Italian mathematicians Cesare Arzelà (1847–1912) and Giulio Ascoli (1843–1896), who both proved versions of it. Variants of this theorem play an important role in mathematical analysis.

The proof is long and difficult and we defer it to Subsection 3.3; we spend the remainder of this subsection seeing how we can use the theorem.

Worked Exercise 3.13

Let

$$A = \{f \in C[0,1] : f(0) = 0 \text{ and } |f(x) - f(y)| \leq |x - y| \text{ for each } x, y \in [0,1]\}.$$

Define $I : C[0,1] \to \mathbb{R}$ by

$$I(f) = \int_0^1 f(x)\,dx.$$

Show that I attains a minimum on A.

Solution

For example, A contains the function $f : [0,1] \to \mathbb{R}$ given by $f(x) = 0$ for $x \in [0,1]$.

We first note that the set A is non-empty. In Book D, Chapter 15, Worked Exercise 5.6, we showed that I is a $(d_{\max}, d^{(1)})$-continuous function. If we can prove that A is a sequentially compact set, then the General Extreme Value Theorem (Theorem 2.13) tells us that I is bounded on A and attains its bounds. In particular, there is an $f_0 \in A$ such that

$$I(f_0) \leq I(f) \quad \text{for all } f \in A,$$

and so I attains its minimum on A.

In fact, f_0 is the function given by $f_0(x) = -x$, but we are not asked to show this.

Thus it only remains for us to show that A is sequentially compact. To do this, we use Theorem 3.12 and verify that:

1. A is $d_{\max}$-closed
2. A is a pointwise bounded family of functions
3. A is an equicontinuous family of functions.

A is closed

Let (f_n) be a sequence of functions in A that $d_{\max}$-converges to $f \in C[0,1]$. To show that A is closed, we must show that $f \in A$.

Since $f_n(0) = 0$ for each $n \in \mathbb{N}$, we have

$$|f(0)| = |f(0) - f_n(0)| \leq d_{\max}(f_n, f).$$

But $d_{\max}(f_n, f)$ is a null sequence. Hence $f(0) = 0$.

For each $x, y \in [0,1]$, we can use the Triangle Inequality (for the modulus) twice and the fact that $f_n \in A$ to estimate that for each $n \in \mathbb{N}$,

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &\leq |f_n(x) - f_n(y)| + 2d_{\max}(f_n, f) \\ &\leq |x - y| + 2d_{\max}(f_n, f). \end{aligned}$$

But $n \in \mathbb{N}$ is arbitrary and $d_{\max}(f_n, f) \to 0$ as $n \to \infty$, so

$$|f(x) - f(y)| \leq |x - y| \quad \text{for each } x, y \in [0,1].$$

Hence f is 1-Lipschitz and $f(0) = 0$, and so $f \in A$, implying that A is closed.

A is a pointwise bounded family of functions

This follows immediately from the result of Exercise 3.7 (with (X, d) equal to $[0, 1]$ with the Euclidean metric, $a = 0$ and $L = 1$).

A is an equicontinuous family of functions

Since each function in A is $(d^{(1)}_{[0,1]}, d^{(1)})$-Lipschitz with Lipschitz constant 1, we can use the result of Worked Exercise 3.9 together with that of Exercise 3.10 to conclude that A is an equicontinuous set of functions.

Since the hypotheses of Theorem 3.12 hold, we conclude that A is sequentially compact, as required.

Exercise 3.12

Let

$$A = \{f \in C[0,1] : f(0) = 0 \text{ and } |f(x) - f(y)| \leq \sqrt{|x-y|} \text{ for all } x, y \in [0,1]\}.$$

Define $F\colon C[0,1] \to \mathbb{R}$ by

$$F(f) = \int_0^1 x f(x)\, dx.$$

(a) Prove that A is $d_{\max}$-sequentially compact.

(b) Prove that F attains a minimum on A.

3.3 Proof of Theorem 3.12

This subsection is not assessed.

In this subsection, we prove Theorem 3.12. The proof is long and involved and you should not be surprised if you find it difficult – your goal should be to understand the outline of the argument and you need not be too concerned with trying to follow every single detail.

For the remainder of this subsection, we suppose that A is a subset of $C[0, 1]$ that has the following properties:

1. A is $d_{\max}$-closed
2. A is pointwise bounded – that is, for each $x \in [0, 1]$, there is an $M_x \geq 0$ such that $|f(x)| \leq M_x$ for each $f \in A$
3. A is equicontinuous – that is, for each $\varepsilon > 0$, there is a $\delta > 0$ such that for each $f \in A$ and each $x, y \in [0, 1]$, $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$.

Our aim is to show that A is $d_{\max}$-sequentially compact – that is, each sequence of functions in A has a $d_{\max}$-convergent subsequence whose limit is also in A.

So let (f_n) be a sequence of functions in A. Our objective is to 'build' a continuous function $f \in C[0,1]$ to which a subsequence of (f_n) $d_{\max}$-converges. We do this in two stages:

1. we find a candidate function f to which a subsequence of (f_n) may converge
2. we show that there is a subsequence of (f_n) that converges to this function f, and conclude that, since A is closed, $f \in A$.

Before embarking on the proof proper, we outline the main ideas.

Sketch of argument

- We start by observing that, since A is pointwise bounded, for each $a \in [0,1]$, there is a real number $M_a \geq 0$ for which $|f_n(a)| \leq M_a$ for each $n \in \mathbb{N}$. In particular, the sequence of real numbers $(f_n(a))$ lies in the $d^{(1)}$-sequentially compact interval $[-M_a, M_a]$ (Corollary 2.2) and so has a convergent subsequence, $(f_{n_k}(a))_{k=1}^{\infty}$ say, that converges to a real number in $[-M_a, M_a]$. Notice, however, that the choice of subsequence depends on the choice of a – different points may require different subsequences, and this is where the difficulty lies.
- But it is possible to find a convergent subsequence of (f_n) that works for two points a and b, say, simultaneously. First, given a, we can find an increasing sequence $1 \leq n_1 < n_2 < \cdots$ of integers so that $(f_{n_k}(a))_{k=1}^{\infty}$ is a convergent sequence converging to a real number, $f(a)$, say. By now considering $(f_{n_k})_{k=1}^{\infty}$ as a sequence of functions in its own right, we can find a subsequence $(n_{k_i})_{i=1}^{\infty}$ of (n_k) so that $(f_{n_{k_i}}(b))$ is also a convergent sequence of real numbers, converging to a real number $f(b)$, say. But $(f_{n_{k_i}}(a))$, being a subsequence of the convergent sequence $(f_{n_k}(a))_{k=1}^{\infty}$, is convergent to $f(a)$, and so we have found a subsequence of the functions (f_n) that converges for both a and b.
- By careful use of induction, we should be able to extend this argument to cope with a countable set of points $\{a_i : i \in \mathbb{N}\}$ (by taking a sequence of subsequences of previously found subsequences).
- Unfortunately, as we saw in Book D, Chapter 16, Theorem 6.7, the interval $[0,1]$ is not countable and so we cannot hope to deal with every point in $[0,1]$ in this way. However, if the function f that we are constructing is to be continuous on $[0,1]$, then we need only find its values on a set that is dense in the interval $[0,1]$, since then the value of f at an arbitrary point a in $[0,1]$ can be found by calculating $\lim_{i\to\infty} f(a_i)$, where (a_i) is a sequence of points lying in the dense set that converges to the point a. This provides us with the key idea needed for proving the theorem: use a *countable* dense subset of $[0,1]$ together with the 'taking subsequences' idea described above to define f on a subset of $[0,1]$, and then show that the f thus defined can be extended to the whole of $[0,1]$ in such a way to have the properties we want – f is continuous, f is a member of A and f is a limit of a subsequence of (f_n).

Book D, Chapter 16, Definition 6.1.

Finding a candidate limit function f

We first identify an appropriate countable dense subset of $[0,1]$ on which we hope to define f via repeated use of subsequences.

Let (a_i) be a sequence of real numbers whose set of values $\{a_i : i \in \mathbb{N}\}$ is dense in $[0,1]$ – that is,

$$\mathrm{Cl}_{(\mathbb{R},d^{(1)})}(\{a_i : i \in \mathbb{N}\}) = [0,1].$$

An example of such a sequence is provided by

$$\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{2}{3}, \tfrac{1}{4}, \tfrac{2}{4}, \tfrac{3}{4}, \ldots, \tfrac{1}{n}, \tfrac{2}{n}, \ldots, \tfrac{n-1}{n}, \tfrac{1}{n+1}, \ldots,$$

This sequence takes every rational value in $(0,1)$ a (countably) infinite number of times.

but there are many other possibilities.

In order to start defining our function f, we try to find a strictly increasing sequence $(n_k)_{k\in\mathbb{N}}$ such that $(f_{n_k}(a_i))_{k=1}^{\infty}$ is a $d^{(1)}$-convergent sequence for each $i \in \mathbb{N}$. This will then allow us to define f on the set of values of the sequence (a_i) by setting $f(a_i) = \lim_{k\to\infty} f_{n_k}(a_i)$ for each $i \in A$.

We construct such a sequence by induction.

We start by choosing a strictly increasing sequence of natural numbers $(n_{1,k})_{k\in\mathbb{N}}$ for which $(f_{n_{1,k}}(a_1))_{k\in\mathbb{N}}$ is a $d^{(1)}$-convergent sequence.

We have shown earlier that we can do this.

Now suppose that for some $j \in \mathbb{N}$, we have a strictly increasing sequence $(n_{j,k})_{k\in\mathbb{N}}$, such that $(f_{n_{j,k}}(a_i))_{k\in\mathbb{N}}$ is a convergent sequence of real numbers for $i = 1, 2, \ldots, j$. Then, since $[-M_{a_{j+1}}, M_{a_{j+1}}]$ is sequentially compact, we can find a subsequence $(n_{j+1,k})_{k\in\mathbb{N}}$ of $(n_{j,k})_{k\in\mathbb{N}}$ such that $(f_{n_{j+1,k}}(a_{j+1}))_{k\in\mathbb{N}}$ is also $d^{(1)}$-convergent.

This subsequence is also strictly increasing.

Proceeding by induction, we find a sequence of strictly increasing sequences $((n_{j,k})_{k\in\mathbb{N}})_{j\in\mathbb{N}}$ for which:

1. $(n_{j+1,k})_{k\in\mathbb{N}}$ is a subsequence of $(n_{j,k})_{k\in\mathbb{N}}$, for each $j \in \mathbb{N}$
2. $n_{k,k} < n_{k+1,k+1}$ for each $k \in \mathbb{N}$, since $n_{k,k} \le n_{k+1,k} < n_{k+1,k+1}$
3. for each $j \in \mathbb{N}$, $(f_{n_{j,k}}(a_i))_{k\in\mathbb{N}}$ is convergent for $i = 1, 2, \ldots, j$.

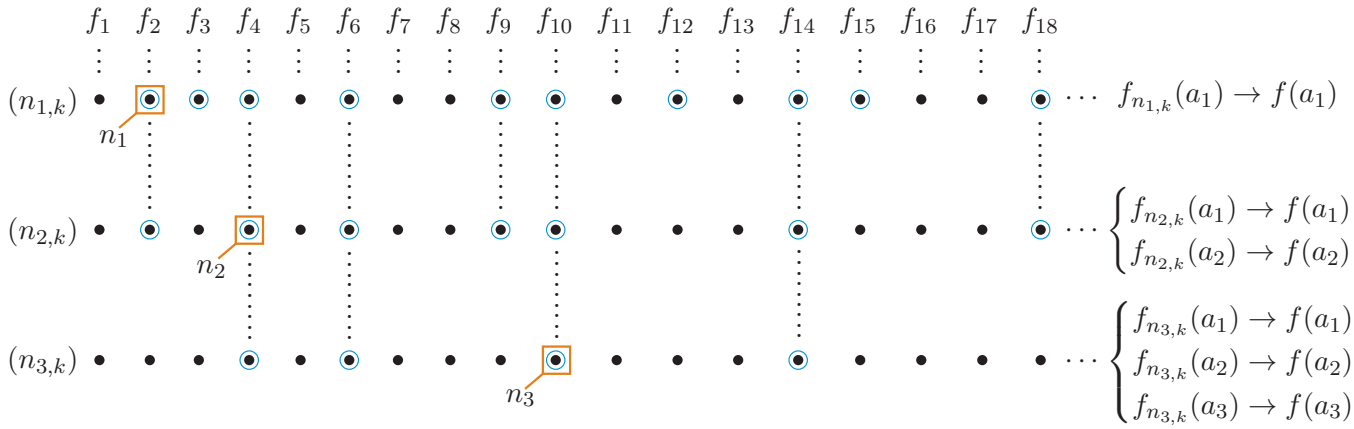


Figure 3.1 This figure shows how we construct the diagonal sequence (n_k), ensuring that the sequence $(f_{n_k}(a_i))_{k=1}^{\infty}$ is convergent for each i

In Book D, Chapter 16, we used a diagonalisation argument to show that $(0,1)$ is uncountable.

We now consider the diagonal sequence $(n_k)_{k=1}^{\infty}$, where $n_k = n_{k,k}$. Then, for each $i \in \mathbb{N}$, $(f_{n_k}(a_i))_{k\in\mathbb{N}}$ is a $d^{(1)}$-convergent sequence, since $(n_k)_{k=i}^{\infty}$ is a subsequence of $(n_{i,k})_{k\in\mathbb{N}}$ and $(f_{n_{i,k}}(a_i))_{k=1}^{\infty}$ is convergent (see Figure 3.1).

We now start to define the function f that will turn out to be the $d_{\max}$-limit of the sequence (f_{n_k}). We begin by defining its value for each a_i in our dense sequence (a_i) by setting

$$f(a_i) = \lim_{k\to\infty} f_{n_k}(a_i), \text{ for each } i \in \mathbb{N}.$$

This function is well defined because if $a_i = a_{i+j}$ for some $j \in \mathbb{N}$, then

$$f(a_i) = \lim_{k\to\infty} f_{n_k}(a_i) = \lim_{k\to\infty} f_{n_k}(a_{i+j}) = f(a_{i+j}).$$

Thus we have defined a function $f\colon \{a_i : i \in \mathbb{N}\} \to \mathbb{R}$.

It remains to extend the definition of f to arbitrary points in $[0,1]$, and this is where we use the fact that $\{a_i : i \in \mathbb{N}\}$ is dense in $[0,1]$. For each $a \in [0,1]$, there is a subsequence $(a_{i_p})_{p=1}^{\infty}$ of $(a_i)_{i=1}^{\infty}$ for which $a_{i_p} \to a$ as $p \to \infty$. We would like to use this subsequence to define $f(a)$ by setting

$$f(a) = \lim_{p\to\infty} f(a_{i_p}).$$

Unfortunately there are two potential problems:

(i) given a subsequence (a_{i_p}) of (a_i) that converges to a, it is not obvious that $(f(a_{i_p}))$ is a convergent sequence

(ii) even if $(f(a_{i_p}))$ is convergent, there may be another choice of subsequence $(a_{j_p})_{p=1}^{\infty}$ of (a_i) that converges to a for which $(f(a_{j_p}))$ does not converge to $\lim_{p\to\infty} f(a_{i_p})$ – this would give us a different inconsistent choice for the definition of $f(a)$.

Fortunately, the fact that A is equicontinuous enables us to show that neither problem can arise.

We first show that

Claim 1

If $(a_{i_p})_{p\in\mathbb{N}}$ and $(a_{j_p})_{p\in\mathbb{N}}$ are subsequences of (a_i) that both converge to a, then $|f(a_{i_p}) - f(a_{j_p})| \to 0$ as $p \to \infty$.

Proof Let $\varepsilon > 0$ be given: we show that there is $P \in \mathbb{N}$ such that for $p > P$, $|f(a_{i_p}) - f(a_{j_p})| < \varepsilon$, implying the claim.

Since A is equicontinuous, we can choose a $\delta > 0$ so that for each $g \in A$ and each $x, y \in [0,1]$,

$$|g(x) - g(y)| < \tfrac{1}{4}\varepsilon \quad \text{whenever } x, y \in [0,1] \text{ satisfy } |x - y| < \delta. \tag{1}$$

Using the Triangle Inequality (for modulus) three times implies that for each $n \in \mathbb{N}$,

$$
\begin{aligned}
&|f(a_{i_p}) - f(a_{j_p})| \\
&= |f(a_{i_p}) - f_n(a_{i_p}) + f_n(a_{i_p}) - f_n(a) + f_n(a) - f_n(a_{j_p}) + f_n(a_{j_p}) - f(a_{j_p})| \\
&\leq |f(a_{i_p}) - f_n(a_{i_p})| + |f_n(a_{i_p}) - f_n(a)| \\
&\qquad + |f_n(a) - f_n(a_{j_p})| + |f_n(a_{j_p}) - f(a_{j_p})|.
\end{aligned}
$$

Since (a_{i_p}) and (a_{j_p}) both converge to a as $p \to \infty$, there is $P \in \mathbb{N}$ such that for $p > P$, both $|a_{i_p} - a| < \delta$ and $|a_{j_p} - a| < \delta$. For $p > P$, we then have by equation (1), for *any* choice of $n \in \mathbb{N}$,

$$|f_n(a_{i_p}) - f_n(a)| < \tfrac{1}{4}\varepsilon \quad \text{and} \quad |f_n(a) - f_n(a_{j_p})| < \tfrac{1}{4}\varepsilon.$$

For each $p \in \mathbb{N}$, $f_{n_k}(a_{i_p}) \to f(a_{i_p})$ and $f_{n_k}(a_{j_p}) \to f(a_{j_p})$ as $k \to \infty$, and so for each $p > P$ there is an $K \in \mathbb{N}$ for which

$$|f(a_{i_p}) - f_{n_K}(a_{i_p})| < \tfrac{1}{4}\varepsilon \quad \text{and} \quad |f_{n_K}(a_{j_p}) - f(a_{j_p})| < \tfrac{1}{4}\varepsilon.$$

Hence, fixing some $p > P$ and choosing a $K \in \mathbb{N}$ so that these inequalities hold, we deduce that

$$|f(a_{i_p}) - f(a_{j_p})| \leq \tfrac{1}{4}\varepsilon + \tfrac{1}{4}\varepsilon + \tfrac{1}{4}\varepsilon + \tfrac{1}{4}\varepsilon = \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we conclude that

$$|f(a_{i_p}) - f(a_{j_p})| \to 0 \text{ as } p \to \infty, \text{ as claimed.} \tag{2}$$

■

We now show that

Claim 2

If $a_{i_p} \to a$ as $p \to \infty$, then $(f(a_{i_p}))_{p\in\mathbb{N}}$ is $d^{(1)}$-convergent.

Proof Suppose that $a_{i_p} \to a$ as $p \to \infty$. We first show that $(f(a_{i_p}))$ is a bounded sequence: by using the Triangle Inequality (for modulus) twice, we find that for each $p, N \in \mathbb{N}$ we have

$$
\begin{aligned}
|f(a_{i_p})| &= |f(a_{i_p}) - f_n(a_{i_p}) + f_n(a_{i_p}) - f_n(a) + f_n(a)| \\
&\leq |f(a_{i_p}) - f_n(a_{i_p})| + |f_n(a_{i_p}) - f_n(a)| + |f_n(a)|.
\end{aligned}
$$

But A is pointwise bounded, and so $(f_n(a))$ is a bounded sequence and there is $M_a > 0$ for which $|f_n(a)| \leq M_a$ for each $n \in \mathbb{N}$. Moreover, A is equicontinuous, and so we can find $\delta > 0$ for which if $x, y \in [0,1]$ with $|x - a| < \delta$, then $|f_n(x) - f_n(y)| < 1$ for each $n \in \mathbb{N}$. Fix $P \in \mathbb{N}$ for which $|a_{i_p} - a| < \delta$ for $p > P$. Then for each $p > P$, we can find n_k so that $|f(a_{i_p}) - f_{n_k}(a_{i_p})| < 1$. Hence for such a $p > P$ and n_k,

$$|f(a_{i_p})| \leq 1 + 1 + M_a = 2 + M_a.$$

Thus for each $p \in \mathbb{N}$,

$$|f(a_{i_p})| \leq \max\{|f(a_{i_p})| : p = 1, 2, \ldots, P\} + 2 + M_a < \infty,$$

and so $(f(a_{i_p}))_{p=1}^{\infty}$ is a bounded sequence of real numbers. Thus by the Monotone Convergence Theorem we can find a convergent subsequence

$(f(a_{i_{p_j}}))_{j=1}^{\infty}$ with limit L, say. It now follows from the Triangle Inequality and equation (2) that

$$|f(a_{i_p}) - L| \leq |f(a_{i_p}) - f(a_{i_{p_j}})| + |f(a_{i_{p_j}}) - L|.$$

The right-hand side of this inequality can be made arbitrarily small by choosing p and j sufficiently large, and so $\lim_{p\to\infty} f(a_{i_p})$ exists (and is L). ■

Claim 1 tells us that if (a_{i_p}) and (a_{j_p}) both converge to a and $(f(a_{i_p}))$ is convergent, then $f((a_{j_p}))$ is also convergent and converges to the same value as $(f(a_{i_p}))$. Hence we can define the value of f at a to be $\lim_{p\to\infty} f(a_{i_p})$, where a_{i_p} is any subsequence of (a_i) that converges to a. This completes our definition of the function $f\colon [0,1] \to \mathbb{R}$.

A subsequence of (f_n) converges to f

It only remains to show that $f \in A$: we need to show that f is continuous (and so in $C[0,1]$) and there is a sequence of functions (f_n) from A that converges to f for $d_{\max}$. We do this by using the notion of *uniform convergence* and show that $f_{n_k} \to f$ uniformly as $k \to \infty$ (where n_k is the strictly increasing sequence we found previously). It then follows from Book D, Chapter 15, Theorem 5.8 that $f \in C[0,1]$, with $f_{n_k} \overset{d_{\max}}{\to} f$ as $k \to \infty$ and so $f \in A$, since A is $d_{\max}$-closed.

Recall from Book D, Chapter 15, Definition 4.4 that a sequence (f_k) of functions $f_k\colon [0,1] \to \mathbb{R}$ **converges uniformly** on $[0,1]$ to the function $f\colon [0,1] \to \mathbb{R}$ if the function $f_k - f$ is bounded for each $k \in \mathbb{N}$, and the sequence (M_k) defined by $M_k = \sup\{|f_k(x) - f(x)| : x \in [0,1]\}$ is a (real) null sequence.

We split the proof that $f_{n_k} \to f$ uniformly into three steps:

(1) we show that (f_{n_k}) converges to f pointwise, that is, for each $a \in [0,1]$, $f_{n_k}(a) \to f(a)$ as $k \to \infty$

(2) we show that f is uniformly continuous and so, by Lemma 3.6, f is continuous

(3) we use steps (1) and (2) to show that $f_{n_k} \to f$ uniformly as $k \to \infty$.

Step (1)

$f_{n_k}(a) \to f(a)$ as $k \to \infty$ for each $a \in [0,1]$.

Let $a \in [0,1]$, and let $\varepsilon > 0$ be given. We must show that there is a $K \in \mathbb{N}$ such that $|f(a) - f_{n_k}(a)| < \varepsilon$ for each $k > K$.

Since A is equicontinuous, we can certainly choose $\delta > 0$ so that for each $x, y \in [0,1]$ if $|x - y| < \delta$, then $|f_n(x) - f_n(y)| < \frac{1}{3}\varepsilon$ for every $n \in \mathbb{N}$. In particular, whenever $x \in [0,1]$ with $|x - a| < \delta$, then $|f_n(x) - f_n(a)| < \frac{1}{3}\varepsilon$ for any $n \in \mathbb{N}$.

Since $(a_i)_{i\in\mathbb{N}}$ is dense in $[0,1]$, there is a subsequence $(a_{i_p})_{p\in\mathbb{N}}$ for which $a_{i_p} \to a$ as $p \to \infty$. And so, by Claim 1, $f(a_{i_p}) \to f(a)$ as $p \to \infty$.

Hence we can find a $P \in \mathbb{N}$ such that for $p > P$, $|a_{i_p} - a| < \delta$ and $|f(a_{i_p}) - f(a)| < \frac{1}{3}\varepsilon$.

On fixing such a $p > P$, using the Triangle Inequality (for modulus) twice gives

$$\begin{aligned}|f(a) - f_{n_k}(a)| &\leq |f(a) - f(a_{i_p})| + |f(a_{i_p}) - f_{n_k}(a_{i_p})| + |f_{n_k}(a_{i_p}) - f_{n_k}(a)| \\ &< \tfrac{1}{3}\varepsilon + |f(a_{i_p}) - f_{n_k}(a_{i_p})| + \tfrac{1}{3}\varepsilon.\end{aligned}$$

But $f_{n_k}(a_{i_p}) \to f(a_{i_p})$ as $k \to \infty$, and so there is $K \in \mathbb{N}$ (that depends on p) so that $|f(a_{i_p}) - f_{n_k}(a_{i_p})| < \frac{1}{3}\varepsilon$ for each $k > K$. Hence for $k > K$,

$$|f(a) - f_{n_k}(a)| < \varepsilon.$$

Since ε is arbitrary, we conclude that $f_{n_k}(a) \to f(a)$ as $k \to \infty$.

Step (2)

f is uniformly continuous (and so is continuous).

Let $\varepsilon > 0$ be given.

Since A is an equicontinuous set, we can find $\delta > 0$ so that, for each $a, b \in [0,1]$ with $|b - a| < \delta$, we have for every $n \in \mathbb{N}$, $|f_n(b) - f_n(a)| < \frac{1}{3}\varepsilon$.

For any $a, b \in [0,1]$, by using the Triangle Inequality (for modulus) twice we find that for each $k \in \mathbb{N}$,

$$\begin{aligned}|f(b) - f(a)| &\leq |f(b) - f_{n_k}(b)| + |f_{n_k}(b) - f_{n_k}(a)| + |f_{n_k}(a) - f(a)| \\ &< |f(b) - f_{n_k}(b)| + \tfrac{1}{3}\varepsilon + |f_{n_k}(a) - f(a)|.\end{aligned}$$

But, by the result of Step (1), $f_{n_k}(b) \to f(b)$ and $f_{n_k}(a) \to f(a)$ as $k \to \infty$. Hence there is a $K \in \mathbb{N}$ so that, for $k > K$, both $|f(b) - f_{n_k}(b)| < \frac{1}{3}\varepsilon$ and $|f_{n_k}(a) - f(a)| < \frac{1}{3}\varepsilon$. Thus on fixing such a $k > K$, we find

$$|f(b) - f(a)| < \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon = \varepsilon.$$

That is, f is uniformly continuous.

Step (3)

$f_{n_k} \to f$ uniformly as $k \to \infty$.

Since f and f_{n_k} are continuous functions, their difference is bounded for each k. It remains to show that (M_k) given by $M_k = \sup\{|f(a) - f_{n_k}(a)| : a \in [0,1]\}$ is a null sequence. To do this, it is enough to show that for each $\varepsilon > 0$, there is a $K \in \mathbb{N}$ so that $|f(a) - f_{n_k}(a)| < \varepsilon$, for each $k > K$ and each $a \in [0,1]$.

Let $\varepsilon > 0$ be given. Since A is equicontinuous and (by Step (2)) f is uniformly continuous, there is $\delta > 0$ so that if $a, b \in [0,1]$ with $|a - b| < \delta$, then both $|f(b) - f(a)| < \frac{1}{3}\varepsilon$ and for each $n \in \mathbb{N}$, $|f_n(b) - f_n(a)| < \frac{1}{3}\varepsilon$.

Let

$$B = \{j\delta : j = 0, 1, 2, \ldots\} \cap [0,1].$$

(So B is a finite set and each $a \in [0,1]$ is at most a distance δ away from a point in B.)

The result of Step (1) tells us that for each $b \in B$, we can find $K_b \in \mathbb{N}$ so that $|f(b) - f_{n_k}(b)| < \frac{1}{3}\varepsilon$ for $k > K_b$.

Let $K = \max\{K_b : b \in B\}$, then K exists and is finite, since B is finite.

If $k > K$, then for each $a \in A$, on choosing $b \in B$ so that $|b - a| < \delta$, we can use the Triangle Inequality (for modulus) twice, to estimate that

$$\begin{aligned} |f(a) - f_{n_k}(a)| &\leq |f(a) - f(b)| + |f(b) - f_{n_k}(b)| + |f_{n_k}(b) - f_{n_k}(a)| \\ &< \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon = \varepsilon. \end{aligned}$$

Hence $f_{n_k} \to f$ uniformly as $k \to \infty$ and we are done.

4 Compact metric spaces

So far we have studied *sequential* compactness. It is natural to wonder why the term 'compactness' is qualified by the word 'sequential' (as mathematicians are not known for using two words when one will suffice). You may have encountered unqualified use of the term 'compactness' in the *History Reader*, and you may have wondered whether the two terms are the same. In this section we define the notion of a *compact* metric space, and show that for a metric space, the two concepts are equivalent.

Compactness is defined in terms of the open sets of a space. We originally defined open sets as being the complements of closed sets but, as with many mathematical concepts, there are alternative characterisations, and the open ball characterisation, given in Book D, Chapter 16, Theorem 2.9 can be more helpful when thinking about compactness. We restate the result here for reference.

Let (X, d) be a metric space and let A be a subset of X. Then A is d-open if, and only if, for each $a \in A$ there is $r > 0$ so that $B_d(a, r) \subseteq A$.

In other words, a d-open set is a set that contains a d-open ball around every one of its points (Figure 4.1). This characterisation can be useful when proving directly that a set is d-open.

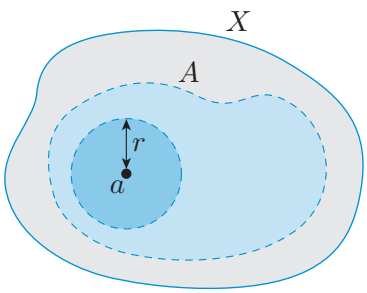


Figure 4.1

We showed that all subsets of a discrete metric space are open in Example 2.5 of Chapter 16, using the fact that an open set is the complement of a closed set.

Worked Exercise 4.1

Use the open set characterisation to prove that the set $\{p\}$, containing the single point $p \in [0,1]$, is d_0-open in $[0,1]$ when $[0,1]$ is considered as a metric space with the discrete metric.

Solution

We need to find an r such that $B_{d_0}(p, r) \subseteq \{p\}$. Take $r = \frac{1}{2}$. Then $B_{d_0}(p, r) = \{x \in [0, 1] \text{ such that } d_0(x, p) < \frac{1}{2}\} = \{p\}$ since $d_0(x, p)$ is equal to 1 for any $x \neq p$. Therefore $B_{d_0}(p, \frac{1}{2}) \subseteq \{p\}$. (In fact, any r with $0 < r \leq 1$ will suffice.)

Exercise 4.1

Use the open set characterisation to show that (a, b) is d-open in $\mathbb{R}$, where d is the metric induced by the usual Euclidean metric on $\mathbb{R}$.

Compactness, as is the case with sequential compactness, is a property of the whole space. Therefore we do not focus our attention on single open sets, but consider collections of open sets that 'cover' the whole space. By this we mean a collection of open sets, $\{U_i : i \in I\}$, which is such that each element of our space is contained within at least one of these open sets. Formally we define an *open cover* as follows.

Definition 4.2 *Open cover*

Let (X, d) be a metric space and let $A \subseteq X$.

A collection $\mathcal{S}$ of d-open subsets of X is an **open cover** of A if

$$A \subseteq \bigcup_{U \in \mathcal{S}} U;$$

that is, for each $a \in A$, there exists $U \in \mathcal{S}$ such that $a \in U$; see Figure 4.2.

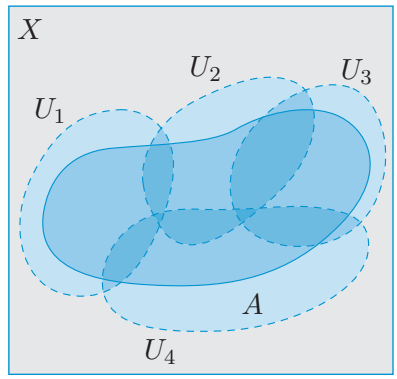


Figure 4.2

Remark

If we wish to emphasise the metric, we say d-open cover.

Let $A = [0, 1]$ with the usual metric, $d^{(1)}$, induced by the Euclidean metric on $\mathbb{R}$. There are many different open covers of A. We begin with the following example.

Worked Exercise 4.3

Show that

$$\mathcal{S}_1 = \{(x - \tfrac{1}{10}, x + \tfrac{1}{10}) : x \in [0, 1]\}$$

is a $d^{(1)}$-open cover of $[0, 1]$, where $d^{(1)}$ is the usual Euclidean metric on $\mathbb{R}$.

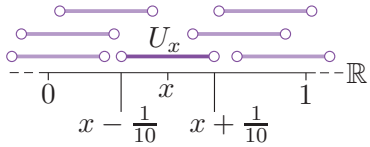


Figure 4.3

Solution

Let $U_x = (x - \frac{1}{10}, x + \frac{1}{10})$. If $x \in [0,1]$, then $x \in U_x$ and so

$$[0,1] \subseteq \bigcup_{x \in [0,1]} U_x = \bigcup_{U_x \in \mathcal{S}_1} U_x \quad \text{(Figure 4.3)}.$$

Also, each set U_x in $\mathcal{S}_1$ is d-open. So $\mathcal{S}_1$ is an open cover of $[0,1]$.

You may have noticed that many of the sets in $\mathcal{S}_1$ overlap, suggesting that $[0,1]$ can be covered by a smaller collection of sets. In fact, $[0,1]$ can be covered by just *finitely* many of the sets in $\mathcal{S}_1$!

Exercise 4.2

Find a finite subcollection of sets in $\mathcal{S}_1$ that covers $[0,1]$.

Now consider

$$\mathcal{S} = \{(0,q) : q \in \mathbb{Q} \cap (0,\infty)\}.$$

This is a collection of $d^{(1)}$-open sets with respect to the Euclidean metric, $d^{(1)}$ on $\mathbb{R}$, but it is not a cover of $[0,1]$, for no set in $\mathcal{S}$ contains the point 0. However,

$$\mathcal{S}_2 = \mathcal{S} \cup \{(-\tfrac{1}{100}, \tfrac{1}{100})\}$$

is an open cover of $[0,1]$.

Exercise 4.3

(a) Show that $\mathcal{S}_2$ is an open cover of $[0,1]$.

(b) Find a finite subcollection of sets in $\mathcal{S}_2$ that covers $[0,1]$.

In fact, it turns out that *any* $d^{(1)}$-open cover of $[0,1]$, contains a finite subcollection of sets that covers $[0,1]$. Before proving this result, we make the following definition.

Definition 4.4 *Finite subcover*

Let (X,d) be a metric space and let $A \subseteq X$. Let $\mathcal{S}$ be a d-open cover of A. A collection $\mathcal{R}$ of d-open subsets of X is a **finite (d-open) subcover** of A from $\mathcal{S}$ if:

(a) $\mathcal{R} \subseteq \mathcal{S}$

(b) $A \subseteq \bigcup_{U \in \mathcal{R}} U$

(c) $\mathcal{R}$ is a finite collection of sets.

Theorem 4.5

Every $d^{(1)}$-open cover of the interval $[0,1]$ has a finite subcover.

Proof Let $\mathcal{S}$ be any $d^{(1)}$-open cover of $[0,1]$. In other words, $\mathcal{S}$ is any collection of $d^{(1)}$-open sets of $\mathbb{R}$ which covers $[0,1]$. We must show that $\mathcal{S}$ contains a finite subcover of $[0,1]$.

Let A be the set of points $x \in [0,1]$ for which the interval $[0,x]$ is covered by a finite subcollection of sets from $\mathcal{S}$.

Considering A is not an obvious thing to do, but it lies at the heart of the proof.

Certainly $0 \in A$, since 0 belongs to at least one set from $\mathcal{S}$ (otherwise $\mathcal{S}$ would not be an open cover of $[0,1]$) and so that set (alone) is a finite subcover of 0. Hence A is not empty.

Observe that if $x \in A$ and $0 \leq y \leq x$, then $y \in A$ since, if $\{U_1, U_2, \ldots, U_n\}$ is a finite subset of $\mathcal{S}$ that covers $[0,x]$, then it also covers $[0,y] \subseteq [0,x]$. Thus A is an interval.

The set A is bounded above by 1, so it must have a least upper bound; denote it by a. Clearly $0 \leq a \leq 1$. If we can show that

The existence of the least upper bound follows from the fact that we are working in $\mathbb{R}$.

$$a \in A \quad \text{and} \quad a = 1,$$

then the proof will be complete.

Proof that $a \in A$

If $a = 0$, then we know that $a \in A$. So suppose that $a \in (0,1]$. There is a set $U \in \mathcal{S}$ with $a \in U$. Since U is open, we can find r with $0 < r < a$ such that $(a - r, a + r) \subseteq U$. Consider now the point $b = a - \frac{1}{2}r \in (0,1]$ (Figure 4.4). This point certainly lies in A – for otherwise, since A is an interval, each point x between b and a is not in A, contradicting the fact that a is the *least* upper bound of A. Thus the interval $[0,b]$ has a finite subcover $\{U_1, U_2, \ldots, U_n\}$ from $\mathcal{S}$. But then $\{U\} \cup \{U_1, U_2, \ldots, U_n\}$ is a finite subcover from $\mathcal{S}$ of $[0,a]$. It follows that $a \in A$.

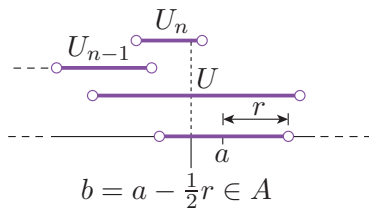


Figure 4.4

Proof that $a = 1$

Suppose, for a contradiction, that $a < 1$. Then, since $a \in A$, the interval $[0,a]$ has a finite subcover $\{U_1, U_2, \ldots, U_n\}$ from $\mathcal{S}$. In particular, $a \in U_i$, for some $i \in \{1, 2, \ldots, n\}$. For this open set U_i, by Book D, Chapter 16, Theorem 2.9 we can find an $r > 0$, with $a + r < 1$ and such that $B_d(a,r) \subseteq U_i$. In particular, $a + \frac{1}{2}r \in U_i$, and so the interval $[0, a + \frac{1}{2}r]$ has a finite subcover from $\mathcal{S}$: namely, $\{U_1, U_2, \ldots, U_n\}$. Thus since $a + \frac{1}{2}r < a + r < 1$, it follows that $a + \frac{1}{2}r \in A$, contradicting the fact that a is an *upper bound* of A. We deduce that $a = 1$.

Note that we use the fact that the U_i are an *open* cover here.

It follows that $\mathcal{S}$ contains a finite subcover of $[0,1]$. ■

Remark

This proof is a classic, created even before formal set theory was invented. An analogous proof can be used to show that any closed bounded interval $[a,b]$ has this property.

Note that it is important that we consider *open* covers of $[0,1]$ – if the cover is not open, there may be no finite subcover.

Exercise 4.4

With respect to the Euclidean metric on $\mathbb{R}$, show that there is an infinite cover of $[0,1]$ by *non-open* sets that has no finite subcover.

We have seen that every open cover of $[0,1]$ has a finite subcover. We now show that the open interval $(0,1)$ does not have this property and ask you to show that the unbounded interval $[0,\infty)$ also does not have this property.

Worked Exercise 4.6

Consider $\mathbb{R}$ with the Euclidean metric and let A be the bounded open set $(0,1)$. Find an open cover of A that has no finite subcover.

Solution

There are many open covers with this property. One such open cover is

$$\mathcal{S} = \{(\tfrac{1}{n}, 1) : n = 2, 3, 4, \ldots\} = \{(\tfrac{1}{2}, 1), (\tfrac{1}{3}, 1), (\tfrac{1}{4}, 1), \ldots\}.$$

For example, if $x = \frac{1}{100}$, then $(\frac{1}{101}, 1)$ contains x.

For each $x \in (0,1)$, the interval $(\frac{1}{n}, 1)$ contains x if $n \in \mathbb{N}$ is chosen so that $\frac{1}{n} < x$. Thus $\mathcal{S}$ is a cover of $(0,1)$. Also, the intervals in $\mathcal{S}$ are open. So $\mathcal{S}$ is an open cover of $(0,1)$. Suppose that $\mathcal{R}$ is a finite subcollection of $\mathcal{S}$. We must show that $\mathcal{R}$ does not cover $(0,1)$. Since $\mathcal{R}$ is finite, we can write

$$\mathcal{R} = \{(\tfrac{1}{n_1}, 1), (\tfrac{1}{n_2}, 1), \ldots, (\tfrac{1}{n_N}, 1)\},$$

where we may assume that $n_1 < n_2 < \cdots < n_N$ (Figure 4.5).

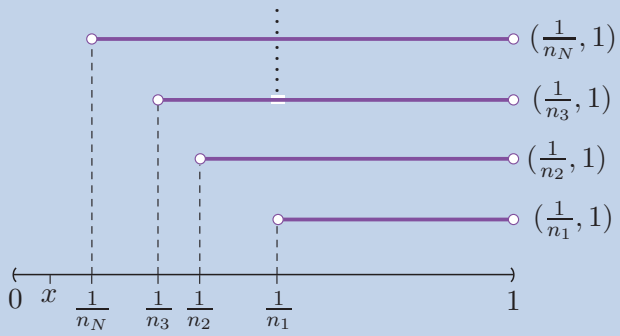


Figure 4.5

Then

$$\bigcup_{U \in \mathcal{R}} U = (\tfrac{1}{n_N}, 1),$$

and so no point $x \in (0, \frac{1}{n_N}]$ is covered. Thus $\mathcal{R}$ is not a cover of $(0, 1)$.

Exercise 4.5

Consider $\mathbb{R}$ with the Euclidean metric and let A be the closed unbounded set $[0, \infty)$. Find an open cover of A that has no finite subcover.

4.1 The Boundedness Theorem revisited

The results of Worked Exercise 4.6 and Exercise 4.5 can be generalised to show that not all $d^{(1)}$-open covers of the open interval (a, b), the half-open intervals $(a, b]$, $[a, b)$, or the unbounded intervals $[a, \infty)$, (a, ∞), $(-\infty, a]$ and $(-\infty, a)$ have finite subcovers. Thus the finite subcover property applies only to intervals that are $d^{(1)}$-closed and bounded.

We now see that the statement 'every open cover of $[0, 1]$ contains a finite subcover' seems to capture some fundamental property of the closed bounded interval $[0, 1]$ that other types of interval fail to possess.

This property enables us to reduce the infinite to the finite. We now give a second proof of the Boundedness Theorem for the interval $[0, 1]$ in which we make use of this ability.

The Boundedness Theorem is Theorem 1.3.

Boundedness Theorem

Let $f : [0, 1] \to \mathbb{R}$ be a continuous function on $[0, 1]$. Then f is bounded on $[0, 1]$. That is, there is a real number M such that $|f(x)| \leq M$ for each $x \in [0, 1]$.

Proof Consider

$$\mathcal{S} = \{U_i : U_i = (i - 1, i + 1),\ i \in \mathbb{Z}\}.$$

Each point $x \in \mathbb{R}$ belongs to an interval U_i. Also, each interval U_i is open. So $\mathcal{S}$ is an open cover of $\mathbb{R}$.

$\mathcal{S}$ is an infinite collection of open sets covering $\mathbb{R}$.

Corresponding to each set U_i is the subset $f^{-1}(U_i)$ of the domain $[0, 1]$. Since f is continuous on $[0, 1]$, the preimage set $f^{-1}(U_i)$ is an open set in $[0, 1]$ (by Book D, Chapter 16, Corollary 7.2). For each $x \in [0, 1]$, $f(x)$ must belong to at least one of the covering sets U_i, so $x \in f^{-1}(U_i)$, and

$$\mathcal{R} = \{f^{-1}(U_i) : i \in \mathbb{Z}\}$$

is an open cover of $[0, 1]$ – that is,

$$[0, 1] \subseteq \bigcup_{i \in \mathbb{Z}} f^{-1}((i - 1, i + 1)).$$

This is where we use compactness to reduce the collection $\mathcal{R} = \{f^{-1}(U_i) : i \in \mathbb{Z}\}$ to the finite collection $\mathcal{R}' = \{f^{-1}(U_i) : -N \leq i \leq N\}$.

Now Theorem 4.5 implies that we can find a finite subcover of $[0, 1]$ from $\mathcal{R}$ – that is, we can find $N \in \mathbb{N}$ such that $[0, 1]$ is covered by the collection $\mathcal{R}' = \{f^{-1}(U_i) : -N \leq i \leq N\}$. (Note that this collection may contain more sets than we need.)

So

$$[0, 1] \subseteq \bigcup_{i=-N}^{N} f^{-1}(U_i).$$

Hence

$$f([0, 1]) \subseteq \bigcup_{i=-N}^{N} U_i = (-N - 1, N + 1).$$

But this means that

$$-N - 1 < f(x) < N + 1 \quad \text{for all } x \in [0, 1].$$

So f is bounded. ■

Thus the finite subcover property of the unit interval is all that is needed to show that a continuous function on that interval is bounded! In particular, note the way in which we used this property to move from an infinite collection of sets, $\bigcup_{i\in\mathbb{Z}} f^{-1}((i - 1, i + 1))$, to a *finite* collection; and that it was this move, from the infinite to the finite, that enabled us to find our bound for the function.

4.2 The definition of a compact metric space

It is also possible to prove the Extreme Value Theorem using the existence of finite subcovers, rather than using the existence of convergent subsequences.

In Subsection 1.1 we saw that every sequence in the closed and bounded interval $[0, 1]$ (with the Euclidean metric) has a convergent subsequence. We used this rather remarkable property in proofs of both the Boundedness Theorem and the Extreme Value Theorem. We also saw, in the discussion after Exercise 1.2, that the unbounded set $\mathbb{R}$ and the open set $(0, 1)$ do not have this property. In Theorem 4.5 we saw that every open cover of the closed and bounded interval $[0, 1]$ (with the Euclidean metric) has a finite subcover. We used this equally remarkable property in a second proof of the Boundedness Theorem (and noted that it could also be used to prove the Extreme Value Theorem). We also saw that unbounded sets, and sets that are not closed, do not possess this property.

In other words, either of these two properties:

1. every sequence has a convergent subsequence
2. every open cover has a finite subcover

can be used to prove results about continuous functions on closed and bounded sets. The first property leads us to a definition of sequential compactness. We now turn our attention to the second, which leads us to the definition of *compactness*.

Definition 4.7

A metric space (X, d) is **compact** if each d-open cover of X contains a finite subcover of X.

A subset $A \subseteq X$ is **compact** if (A, d_A) is compact as a metric space in its own right.

As usual, d_A is the metric on A induced by the metric d on X.

Remark

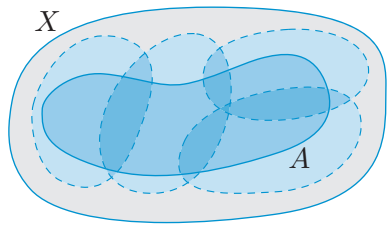
If we want to emphasise the metric, then we refer to a compact space (X, d) as a d-compact metric space.

We now pause to unpack exactly what this definition means when we are working with subsets, and to check the consistency of our definitions.

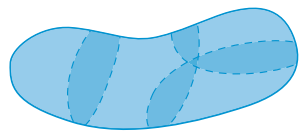
Let (X, d) be a metric space with $A \subseteq X$. In Definition 4.4 we defined a d-open cover of A as being a collection of d-open subsets of X which 'cover' A.

There is no problem with applying this definition to the whole set X. We simply take a collection of d-open subsets of X which cover all of X. We do need to be a little careful when we apply this definition to a compact *subset* of X.

Definition 4.7 states that A is a compact subset of X if (A, d_A) is compact as a metric space in its own right. That is, A is a compact subset if every d_A-open cover of A has a finite subcover. We know, from Theorem 2.1 of Chapter 21, that a set $U \subseteq A$ is d_A-open in A if, and only if, there is a d-open set $V \subseteq X$ with $U = A \cap V$. Therefore, for every d-open cover of A, by d-open sets, $V_i \subseteq X$, we have a corresponding d_A-open cover of A, given by the collection of all sets of the form $V_i \cap A$. Hence, if an open cover of A by d-open sets in X has a finite subcover, then the corresponding open cover by d_A-open sets in A also has a finite subcover, and vice versa; see Figure 4.6. Therefore our definitions are consistent.



V_i covering $A \subseteq X$



$V_i \cap A$ covering A

Figure 4.6

Corollary 4.8

The interval $[0, 1]$, with the metric induced by the usual Euclidean metric on $\mathbb{R}$, is compact.

Proof This follows from Theorem 4.5. ■

A similar argument can be used to show that each $d^{(1)}$-closed and bounded interval $[a, b]$ is compact.

$(0, 1)$ and $[0, \infty)$ are not compact, as is the case for any interval that is not both closed and bounded.

We state this formally in Theorem 5.14

Remark

Warning: when trying to prove that a given space X is compact, it is tempting, but wrong, to choose some *particular* open cover and demonstrate that it has a finite subcover. This does not prove that X is

compact. We must show that *every* open cover has a finite subcover. Thus a proof of compactness typically begins by considering an *arbitrary* open cover $\mathcal{S}$ of X. To show that a space is not compact it is, of course, sufficient to find a single open cover which does not have a finite subcover.

4.3 Examples of compact spaces

We now ask you to examine some metric spaces to see which are compact.

Exercise 4.6

Let (X, d) be a metric space. Show that $\emptyset$ is a compact subset of X.

The most important examples for us to examine for compactness are subsets of $\mathbb{R}^n$. Ideally we would like to be able to characterise these completely – that is we would like necessary and sufficient conditions for a subset of $\mathbb{R}^n$ to be compact. We already know, from Proposition 2.11, that any $d^{(n)}$-closed and bounded subset of $\mathbb{R}^n$ (with the usual Euclidean metric) is sequentially compact. We would like to use this fact to help us to characterise compact subsets of $\mathbb{R}^n$, so we defer discussion of these subsets until Section 5, where we show that any sequentially compact metric space is compact.

For the moment, we first consider finite sets. It is no surprise to discover that these are compact.

Exercise 4.7

Let (X, d) be a metric space. Show that each finite subset of X is compact.

We now turn our attention to the discrete metric.

Example 4.9

Here every subset of X is open.

Consider (X, d_0), where d_0 is the discrete metric. We know from Exercise 4.7 that any finite set is compact, so suppose that X is an infinite set and let A be an infinite subset of X. In this case, A is not compact. To show this, it is enough to exhibit one open cover of A that has no finite subcover: the collection $\{\{a\} : a \in A\}$ is such a cover.

Exercise 4.8

Show that $(\mathbb{R}, d^{(1)})$ is not compact.

5 Sequentially compact and compact are equivalent for metric spaces

We divide the proof that sequential compactness and compactness are equivalent notions for metric spaces into two parts. First we show in Theorem 5.10 that any sequentially compact metric space is compact, and then in Theorem 5.11 we show that compact metric spaces are sequentially compact.

The main idea of the proof of Theorem 5.10 is to show that all the points in a sequentially compact space are 'near' one of the points in a finite set of points. This allows us to 'replace' the space by a finite collection of open sets. As with many proofs preceded by a discussion where some of the words are held within quotation marks, the proof that a sequentially compact metric space is compact is rather long.

We want to show that all the points in a sequentially compact metric space, (X, d), can be approximated by a finite set of points. By this we mean that for each $\varepsilon > 0$, it is possible to find a finite set $N(\varepsilon)$ of points, such that for each point $x \in X$, we can find a $p_x \in N(\varepsilon)$, with $d(x, p_x) < \varepsilon$. This allows us to 'replace' the whole space X by the finite set of points, $N(\varepsilon)$, with an error of at most ε.

The set $N(\varepsilon)$ will depend on ε and the point p_x will also depend on x.

Definition 5.1 *ε-net*

Let (X, d) be a metric space. For $\varepsilon > 0$, an **ε-net** for X is a set $N(\varepsilon) \subseteq X$ such that

$$X \subseteq \bigcup_{p \in N(\varepsilon)} B_d(p, \varepsilon).$$

Remark

When we want to emphasise the metric, we use the notation $N_d(\varepsilon)$.

An ε-net for a metric space (X, d) is a set of points, $\{p_i : i \in I\}$, such that taking a d-open ball, of radius ε, around each of these points, gives us a collection of balls that covers the space X (see Figure 5.1 overleaf).

Remarks

1. Note that an ε-net is a set of *points*. It does not look like a fishing net (or for that matter a net of any other description). It is a (possibly uncountable) set of points such that any other point in the space is 'close' to one of the points in the net.
2. An ε-net provides an approximation to within a distance ε of any point in the space in the following sense. If $N(\varepsilon) = \{p_i : i \in I\}$ is an ε-net for the metric space (X, d), then for each $x \in X$, we can find a $p_i \in N(\varepsilon)$ such that $d(x, p_i) < \varepsilon$. For example, $\mathbb{Z}$ is an ε-net for $\mathbb{R}$ with the Euclidean metric, provided that $\varepsilon > \frac{1}{2}$.

3. For each $\varepsilon > 0$, X is an ε-net for itself.
4. If S is an ε-net for X, then it is also a δ-net for X for each $\delta > \varepsilon$.

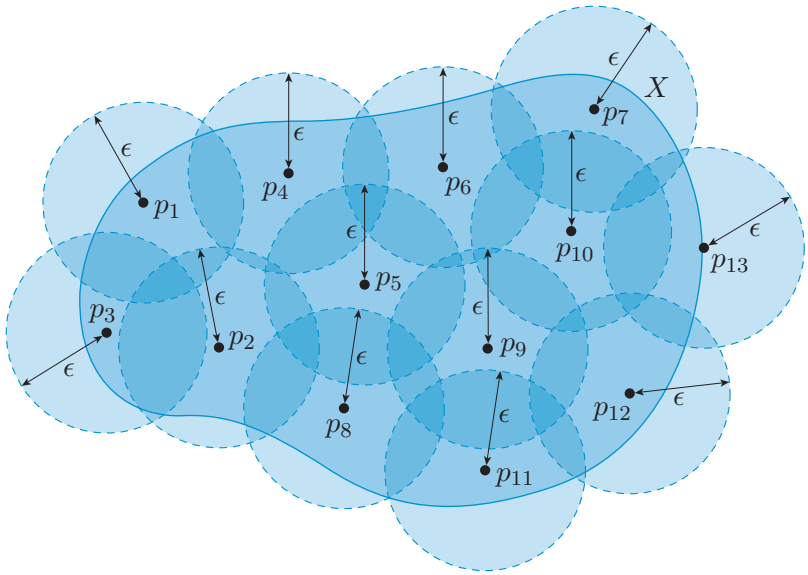


Figure 5.1

Every metric space, (X, d), has at least one ε-net for any ε, namely X itself. (Of course, if X is infinite, then this will be an infinite ε-net.) As the first step of our journey to showing that every sequentially compact space is compact, we show that a sequentially compact space has a *finite* ε-net for each $\varepsilon > 0$. This will not, of course, show that every open cover has a finite subcover; but it will give us a finite collection of sets of radius ε with which to work.

Lemma 5.2

Let (X, d) be a sequentially compact metric space. Then X has a *finite* ε-net for each $\varepsilon > 0$.

Proof Clearly the empty set has a finite ε-net for any ε, so we may assume that $X \neq \emptyset$. Suppose that X is sequentially compact, but that there is an $\varepsilon > 0$ for which there is no finite ε-net for X. We obtain a contradiction by constructing a sequence in X that has no convergent subsequence. We will do this by taking an arbitrary point, $a_1 \in X$, and then taking a second point which is 'more than ε distant' from a_1, then finding a third which is 'more than ε distant' from each of these, and so on, which gives us a sequence where all of the points are spaced at least ε apart (and so it cannot have a convergent subsequence).

Let $a_1 \in X$. Since there is no finite ε-net, $\{a_1\}$ cannot be an ε-net for X. Hence there is an $a_2 \in X$ such that $d(a_1, a_2) \geq \varepsilon$.

Similarly, $\{a_1, a_2\}$ cannot be an ε-net for X, and so there is an $a_3 \in X$ whose distance is at least ε from both a_1 and a_2 – that is,

$$\min\{d(a_1, a_3), d(a_2, a_3)\} \geq \varepsilon.$$

Thus for $\{a_1, a_2, a_3\}$, we have $d(a_i, a_j) \geq \varepsilon$ for each $i \neq j$.

We continue in this way, constructing subsets $\{a_1, a_2, \ldots, a_k\}$ of X that are not ε-nets for X and for which $d(a_i, a_j) \geq \varepsilon$ for $i \neq j$.

Thus we construct a sequence (a_n) such that $d(a_i, a_j) \geq \varepsilon$ for any two distinct terms a_i and a_j in the sequence. In particular, the terms of this sequence do not get arbitrarily close (they are always at least ε apart), and so this sequence has no convergent subsequences, contradicting our assumption that X is sequentially compact. ■

Exercise 5.1

Use Lemma 5.2 to show that a sequentially compact metric space is bounded.

We have proved the following.

Corollary 5.3

A sequentially compact metric space is bounded.

Proof This is Exercise 5.1. ■

We have now seen that a sequentially compact metric space has a finite ε-net, and that this fact can be used to show that any sequentially compact metric space is bounded. However, notice that we can prove that a sequentially compact space is bounded by assuming that a finite ε-net exists *for just one value of* ε. Lemma 5.2 shows something far stronger: that, in a sequentially compact metric space, we can find a finite ε-net for *any* $\varepsilon > 0$.

5.1 Totally bounded sets

We want to show that sequential compactness is equivalent to compactness for metric spaces. We have just shown, in Corollary 5.3, that all sequentially compact metric spaces are bounded. This may lead you to wonder if sequential compactness is equivalent to boundedness. This turns out to be far from the case. In fact, the next few exercises show that, when investigated using only convergent sequences, all metric spaces 'look bounded'! To do this we introduce a new metric, called the **min metric**: $d_{\min}$.

Definition 5.4 *The min metric*

Let (X, d) be a metric space and define a distance function $d_{\min}\colon X \times X \to \mathbb{R}$ as follows:

$$d_{\min}(a,b) = \begin{cases} d(a,b), & \text{if } d(a,b) \leq 1, \\ 1, & \text{if } d(a,b) > 1. \end{cases}$$

First we ask you to check that $d_{\min}$ really is a metric (assuming that d is a metric to start with, of course).

Exercise 5.2

Show that $d_{\min}$ is a metric.

The following may help you to develop a feel for a space with the min metric. Consider the real line, with the metric $d^{(1)}_{\min}$. Imagine that we are standing in this space, at a fixed point, and looking around. Intuitively, when we 'look at' our immediate surroundings, we can see some detail (as points that are 'within 1 of us' retain their original spacings), but once we look further into the distance all of the points appear equally far away. We can easily bound our space by enclosing it within a ball of any radius greater than 1.

We now show that the convergent sequences in any metric space (X, d) are the same as the convergent sequences in the space with the corresponding min metric, $d_{\min}$.

Exercise 5.3

Let (X, d) be a metric space and let (a_n) be a sequence in X. Show that (a_n) is d-convergent if, and only if, (a_n) is $d_{\min}$-convergent.

Exercise 5.3 shows that the convergent sequences for any metric space (X, d) are the same as those for the metric space $(X, d_{\min})$. We now show that any subset of a space with the min metric is $d_{\min}$-bounded.

Exercise 5.4

Let (X, d) be any metric space and let $A \subseteq X$. Show that A is $d_{\min}$-bounded, where $d_{\min}$ is the metric derived from d as described in Definition 5.4.

In particular, the whole space $(X, d_{\min})$ is bounded. Therefore the convergent sequences in any metric space (X, d) are the same as the convergence sequences in the metric space $(X, d_{\min})$, and in this space every set is bounded. So, if we only use convergent sequences to examine a metric space, then any space (X, d) will 'look like' the bounded space

$(X, d_{\min})$. In other words, if we only use sequences to 'look with', then we cannot 'see' unboundedness. Since the definition of sequential compactness only relies on the convergence of certain (sub)sequences, sequential compactness is a property that we need to use sequences to 'see'. Therefore boundedness is not going to be a helpful quality when trying to prove that sequentially compact spaces are compact and we need a stronger concept than boundedness. This stronger concept is that of *total boundedness*, the definition of which is motivated by Lemma 5.2.

Definition 5.5 *Totally bounded*

A metric space (X, d) is **totally bounded** if there is a finite ε-net for each $\varepsilon > 0$.

In a bounded space, we can always find an $\varepsilon > 0$ for which there is a finite ε-net. In a totally bounded space, rather than being satisfied by finding a single ε, and corresponding finite ε-net, we require that we can start with an arbitrary $\varepsilon > 0$ and, no matter how small this ε is, guarantee that we can always find a finite ε-net for this ε. In general, as ε gets smaller, the number of points in the ε-net will get larger, but for a totally bounded space this number must remain finite, however small ε becomes. In the next exercise we ask you to show that the closed and bounded interval $[a, b]$ is, in fact, closed and totally bounded.

Exercise 5.5

Show that $[a, b]$ (with the Euclidean metric) is totally bounded.

We now return very briefly to the min metric. Consider $(\mathbb{R}, d_{\min}^{(1)})$, which we know is bounded. For this space to be totally bounded, it is not sufficient to fix a point and simply enclose our space within a finite-radius ball, where the radius can be as big as we like; we need to be able to find finitely many balls for any given radius ε. Taking $\varepsilon < 1$ means that we are unable to take advantage of the 'contracting effect' of the min metric and so would need infinitely many balls to cover our space. So we see that total boundedness is a stronger concept than boundedness.

In the next exercise, we ask you to show that every compact metric space is totally bounded.

Exercise 5.6

Show directly that each compact metric space (X, d) is totally bounded.

5.2 Sequentially compact metric spaces are compact

In our quest to prove that sequential compactness and compactness are equivalent for metric spaces, we now have the following two results.

Lemma 5.6

Each compact metric space (X, d) is totally bounded.

Proof This is Exercise 5.6. ■

Corollary 5.7

Each sequentially compact metric space is totally bounded.

Proof This follows directly from Lemma 5.2 and Definition 5.5. ■

The situation is looking hopeful. We now show that, for any open cover $\mathcal{S}$ of a sequentially compact space (X, d), we can find a number ε such that the d-open ball of radius ε around each point lies entirely within one of the sets in $\mathcal{S}$. This may seem a rather dull and technical result, but it turns out to be the key to proving that sequentially compact metric spaces are compact.

Lemma 5.8

Let (X, d) be a sequentially compact metric space and let $\mathcal{S}$ be an open cover of X. Then there is an $\varepsilon > 0$ such that, for each $x \in X$, there is a $U \in \mathcal{S}$ with $B_d(x, \varepsilon) \subseteq U$.

Proof Let (X, d) be a sequentially compact metric space, and suppose that the result is false.

In this case, there is an open cover $\mathcal{S}$ of X such that we cannot find an $\varepsilon > 0$ such that, for each $x \in X$, there is a $U \in \mathcal{S}$ with $B_d(x, \varepsilon) \subseteq U$. Therefore for each $\varepsilon > 0$, there must be an $x \in X$ with

This says that the ball $B_d(x, \varepsilon)$ is never entirely contained within an open set of the cover.

$$B_d(x, \varepsilon) - U \neq \emptyset \quad \text{for each } U \in \mathcal{S}.$$

We now use this fact to construct a sequence.

We are taking $\varepsilon = \frac{1}{n}$ here.

For each $n \in \mathbb{N}$, there is an $a_n \in X$ such that

$$B_d(a_n, \tfrac{1}{n}) - U \neq \emptyset \quad \text{for each } U \in \mathcal{S}.$$

Since X is sequentially compact, the sequence $(a_n)_{n=1}^{\infty}$ has a convergent subsequence $(a_{n_k})_{k=1}^{\infty}$ with limit $a \in X$, say.

Since $\mathcal{S}$ is a cover of X, there is a set $U \in \mathcal{S}$ with $a \in U$. Since U is open, there is an $M \in \mathbb{N}$ such that

$$B_d\left(a, \tfrac{2}{M}\right) \subseteq U.$$

Now, $a_{n_k} \to a$ as $k \to \infty$, and so there is a $K \in \mathbb{N}$ such that, for all $k > K$,

$$d(a_{n_k}, a) < \tfrac{1}{M}, \quad \text{and so } a_{n_k} \in B_d\left(a, \tfrac{1}{M}\right).$$

We now recall that, by the definition of the sequence (a_n),

$$B_d\left(a_{n_k}, \tfrac{1}{n_k}\right) - U \neq \emptyset \quad \text{for each } k \in \mathbb{N}.$$

However, if k is chosen so that both $k > K$ and $n_k > M$, then

$$B_d\left(a_{n_k}, \tfrac{1}{n_k}\right) \subseteq B_d\left(a_{n_k}, \tfrac{1}{M}\right) \subseteq B_d\left(a, \tfrac{2}{M}\right) \subseteq U.$$

See Figure 4.1.

This implies that $B_d\left(a_{n_k}, \tfrac{1}{n_k}\right) - U = \emptyset$: a contradiction.

Hence there is an $\varepsilon > 0$ such that, for each $x \in X$, there is a $U \in \mathcal{S}$ with $B_d(x, \varepsilon) \subseteq U$. ■

Definition 5.9 *Lebesgue number*

The number ε given in Lemma 5.8 is called a **Lebesgue number** of the open cover $\mathcal{S}$.

Henri Lebesgue (1875–1941) is primarily remembered by mathematicians for his fundamental contributions to the theory of integration.

We can now prove that sequentially compact metric spaces are compact.

Theorem 5.10

Every sequentially compact metric space is compact.

Proof Let (X, d) be a sequentially compact metric space. We must show that each open cover of X has a finite subcover.

So let $\mathcal{S}$ be an open cover of X. Lemma 5.8 implies that there is an $\varepsilon > 0$ such that, for each $x \in X$, there is a $U \in \mathcal{S}$ with $B_d(x, \varepsilon) \subseteq U$.

Since (X, d) is sequentially compact, Corollary 5.7 implies that (X, d) is totally bounded. Thus there are a finite number of points $x_1, x_2, \ldots, x_N \in X$ such that, for this choice of ε,

$$X \subseteq \bigcup_{i=1}^{N} B_d(x_i, \varepsilon).$$

But, for each x_i, there is a $U_i \in \mathcal{S}$ for which $B_d(x_i, \varepsilon) \subseteq U_i$. Therefore

$$X \subseteq \bigcup_{i=1}^{N} B_d(x_i, \varepsilon) \subseteq \bigcup_{i=1}^{N} U_i.$$

Thus the collection $\{U_1, U_2, \ldots, U_N\}$ is a finite subcover of X from $\mathcal{S}$.

Since $\mathcal{S}$ is an arbitrary open cover of X, the result follows. ■

5.3 Compact metric spaces are sequentially compact

In Theorem 5.10 we proved that all sequentially compact metric spaces are compact. We now show that all compact metric spaces are sequentially compact, thus proving that, for metric spaces, the definitions are equivalent.

Theorem 5.11

Let (X, d) be a compact metric space. Then each sequence in X has a convergent subsequence and so (X, d) is sequentially compact.

Proof The proof is by contradiction, twice.

Suppose that (X, d) is a compact metric space for which there is a sequence (a_n) in X with no convergent subsequences.

We claim that, under this supposition, for each $x \in X$, there are an $\varepsilon_x > 0$ and an $N_x \in \mathbb{N}$ such that

$$a_n \notin B_d(x, \varepsilon_x) \quad \text{for all } n > N_x.$$

We prove this claim by contradiction. Suppose that the claim is not true. Then there is an $x \in X$ such that, for each $\varepsilon > 0$ and each $N \in \mathbb{N}$, there is some $n > N$ for which

$$a_n \in B_d(x, \varepsilon).$$

In particular, for $\varepsilon_1 = \frac{1}{2}$, there is an integer n_1 such that $n_1 > 1$ and

$$a_{n_1} \in B_d(x, \varepsilon_1) = B_d(x, \tfrac{1}{2}).$$

Similarly, for $\varepsilon_2 = \frac{1}{4}$, there is an integer $n_2 > n_1$ such that

$$a_{n_2} \in B_d(x, \varepsilon_2) = B_d(x, \tfrac{1}{4}).$$

Proceeding in this way, we find that for $\varepsilon_k = 2^{-k}$ there is an integer n_k such that $n_k > n_{k-1} > \cdots > n_1$ and

$$a_{n_k} \in B_d(x, \varepsilon_k) = B_d(x, 2^{-k}).$$

This follows from the Squeeze Rule as (2^{-k}) is a null sequence.

Therefore $d(a_{n_k}, x) < 2^{-k}$ and so $(d(a_{n_k}, x))$ is a null sequence and we have a subsequence (a_{n_k}) that converges to x. But we are assuming that no such subsequence exists – a contradiction!

Thus, under the supposition that the sequence (a_n) has no convergent subsequences, we have shown that, for each $x \in X$, there are an $\varepsilon_x > 0$ and an $N_x \in \mathbb{N}$ with

$$B_d(x, \varepsilon_x) \cap \{a_n : n > N_x\} = \emptyset.$$

Clearly the collection

$$\{B_d(x, \varepsilon_x) : x \in X\}$$

This is where we use compactness.

is an open cover of X. Since X is compact, the definition of compactness tells us that there exist $x_1, x_2, \ldots, x_k \in X$ such that

$$X \subseteq \bigcup_{i=1}^{k} B_d(x_i, \varepsilon_{x_i}).$$

In particular, the set $\{a_n : n \in \mathbb{N}\}$ of the sequence is contained in this finite union of balls:

$$\{a_n : n \in \mathbb{N}\} \subseteq \bigcup_{i=1}^{k} B_d(x_i, \varepsilon_{x_i}).$$

By the definition of N_{x_i},

$$B_d(x_i, \varepsilon_{x_i}) \cap \{a_n : n > N_{x_i}\} = \emptyset.$$

Now let $N = \max\{N_{x_1}, N_{x_2}, \ldots, N_{x_k}\}$. It follows that

$$a_n \notin \bigcup_{i=1}^{k} B_d(x_i, \varepsilon_{x_i}) \quad \text{for } n > N$$

– a contradiction, since the whole space is contained in this union!

We conclude that there is no sequence in X without a convergent subsequence. ■

Thus we have:

Theorem 5.12

A metric space is compact if, and only if, it is sequentially compact.

Proof This follows from Theorem 5.11 and Theorem 5.10. ■

Remark

Note that if (X, d) is a metric space and (a_n) lies in a compact subset A of X, then (a_n) has a convergent subsequence with limit in A.

We now know that, for metric spaces, the notions of compact and sequentially compact are equivalent. Therefore our results for sequentially compact spaces hold for compact spaces and we have the following results.

Corollary 5.13

1. Let $A \subseteq X$ where (X, d) is a d-compact metric space. Then A is d-closed if, and only if, A is d_A-compact.
2. Let (X, d) and (Y, d') be metric spaces with X being d-compact. Let $f\colon X \to Y$ be a (d, d')-continuous function. Then the image, $f(X)$, is d'-compact.
3. Let (X, d) and (Y, d') be two compact metric spaces. Then $(X \times Y, e)$ is compact (where e is a product metric).
4. Compactness is a topological invariant.

Proof

1. This follows from Theorem 2.5 and Theorem 5.12.
2. This follows from Lemma 2.6 and Theorem 5.12.
3. This follows from Theorem 2.8 and Theorem 5.12.
4. This follows from Corollary 2.7 and Theorem 5.12.

■

In particular, we have the following version of the Heine–Borel Theorem.

Theorem 5.14 *The Heine–Borel Theorem*

A subset of $\mathbb{R}^n$ is $d^{(n)}$-compact if, and only if, it is $d^{(n)}$-closed and bounded.

Proof This follows from Theorem 2.12 and Theorem 5.12. ■

6 Unions and intersections of compact spaces

Since compactness is expressed in terms of open sets, it can be easier to use compactness (as opposed to sequential compactness) to prove results concerning set operations.

Theorem 6.1

Let (X, d) be a metric space. Then the intersection of any collection of compact subsets of X is compact.

Proof Let $\mathcal{F}$ be a collection of d-compact subsets of X. We must show that $\bigcap_{K\in\mathcal{F}} K$ is d-compact. Theorem 5.12 implies that K is d-sequentially compact and so, by the result of Exercise 2.1, K is d-closed. It follows that $\bigcap_{K\in\mathcal{F}} K$ is a d_K-closed subset (for the induced metric) of any one of the compact sets K in the collection $\mathcal{F}$, and so, by Corollary 5.13, $\bigcap_{K\in\mathcal{F}} K$ is compact. ■

Corollary 6.2

Let (X, d) be a metric space. Then the intersection of any collection of sequentially compact subsets of X is sequentially compact.

Proof This follows from Theorem 6.1 and Theorem 5.12. ■

In Chapter 24, we will investigate some basic properties of fractal sets. You will see that many such sets can be defined as intersections of compact subsets of metric spaces. Theorem 6.1 guarantees that the resulting intersection is compact. However, the intersection can be empty, particularly if a collection $\mathcal{F}$ of compact sets is large. For example, if $\mathcal{F} = \{\{x\} : x \in [0,1]\}$, then $\bigcap_{K\in\mathcal{F}} K = \emptyset$. Fortunately, there is one situation, useful in the study of fractals, where the intersection of a collection of compact sets must be non-empty in a metric space.

Theorem 6.3

Let (X, d) be a metric space, and let K_n $(n = 1, 2, 3, \ldots)$ be non-empty compact sets in X for which

$$K_1 \supseteq K_2 \supseteq K_3 \supseteq \cdots \supseteq K_n \supseteq K_{n+1} \supseteq \cdots .$$

Then $\bigcap_{n=1}^{\infty} K_n$ is a non-empty compact set.

Proof Theorem 6.1 implies that $\bigcap_{n=1}^{\infty} K_n$ is compact. It remains to show that the intersection is non-empty.

Since K_i is compact, for each i, it follows from Corollary 5.13 that each K_n is closed. Thus K_n^c is open, for each n. Moreover, by De Morgan's Second Law,

Chapter 16, Theorem 2.12.

$$\bigcup_{n=1}^{\infty} K_n^c = \left(\bigcap_{n=1}^{\infty} K_n\right)^c .$$

We assume that $\bigcap_{n=1}^{\infty} K_n = \emptyset$ and use proof by contradiction. Then $\{K_n^c : n \in \mathbb{N}\}$ is an open cover of X, and so (in particular) is an open cover of K_1. Since K_1 is compact, there is a number $N \in \mathbb{N}$ for which

$$K_1 \subseteq \bigcup_{n=1}^{N} K_n^c \quad \text{(Figure 6.1)}.$$

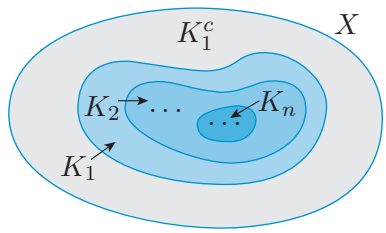


Figure 6.1

Using De Morgan's Second Law again, we see that

$$\bigcup_{n=1}^{N} K_n^c = \left(\bigcap_{n=1}^{N} K_n\right)^c = K_N^c.$$

Therefore $K_1 \subseteq K_N^c$. Hence

$$K_N \subseteq K_1 \subseteq K_N^c.$$

We deduce that K_N is empty – a contradiction. Hence $\bigcap_{n=1}^{\infty} K_n \neq \emptyset$, as required. ■

Proposition 6.4

The finite union of a collection of compact subsets is compact.

Proof We ask you to prove this in Exercise 6.1. ■

Corollary 6.5

The finite union of a collection of sequentially compact subsets is sequentially compact.

Proof This follows from Proposition 6.4 and Theorem 5.12. ■

Exercise 6.1

Let (X, d) be a metric space, and let K_m $(m = 1, \ldots, n)$ be non-empty compact sets in X. Show that the finite union, $\bigcup_{m=1}^{n} K_m$ is a non-empty compact set.

6.1 Epilogue

This section is not assessed.

You may be wondering why we have made a point of adding 'for a metric space' to our results and comments about the equivalence of the definitions of sequential compactness and compactness. This is because the two definitions can be made to apply to all topological spaces, not just to topological spaces which are also metric spaces (metrisable topological spaces) and, for a topological space, the two definitions are not equivalent. A metric space, (X, d), can be viewed as a topological space by taking the d-open sets as its topology, but it is also possible to specify topologies which do not arise from any metric. In a metrisable topological space (one where the open sets arise as the d-open sets for metric d), the notions of sequential compactness and compactness coincide; but in a general topological space this is not true. For example, it turns out that there is a topological space whose underlying set consists of uncountably many copies of the two-point space $\{0, 1\}$ that is compact but not sequentially compact; and there are topological spaces which are sequentially compact but not compact. However, a discussion of these spaces, and the way in which we can define topologies on them, is beyond the scope of this module.

A topological space was defined in Book D, Chapter 16 and that the structure on a topological space is given by specifying a collection of sets, which we call 'open'. This collection is called a *topology*.

Learning outcomes

We list those topics on which we may set assessment questions to test your understanding of the chapter.

After working through this chapter you should be able to:

(a) Explain what it means for a metric space to be sequentially compact.

(b) Identify sequentially compact subsets of many of our standard examples of metric spaces.

(c) Use the Heine–Borel Theorem to identify sequentially compact subsets of Euclidean spaces.

(d) Use the General Extreme Value Theorem.

(e) Know what it means for a collection of functions to be pointwise bounded or equicontinuous.

(f) Use the Arzelà–Ascoli Theorem to identify sequentially compact subsets of $C[0, 1]$.

(g) Explain what it means for a metric space to be compact.

(h) Identify compact subsets of many metric spaces.

(i) Know that compactness and sequential compactness are equivalent in metric spaces.

(j) Know how compact and sequentially compact sets behave under unions and intersections.

Solutions to exercises

Solution to Exercise 1.1

In an arbitrary metric space we have a notion of distance but we do not necessarily have a notion of *order*. In other words, $a > b$ does not automatically make sense. For example, the complex numbers $\mathbb{C}$ with the metric given by $d(w, z) = |w - z|$ where $|a + bi| = \sqrt{a^2 + b^2}$ form a metric space, but it is not obvious which of, for example, $2 + i$ and $1 + 2i$ is, or should be, 'bigger'.

Solution to Exercise 1.2

Points in $C[0, 1]$ are functions, so a sequence in $C[0, 1]$ will be an ordered list of functions, $(f_1, f_2, f_3, \ldots)$, where, for each $n \in \mathbb{N}$, f_n is a continuous function on $[0, 1]$. Therefore the following is an example of a sequence in $C[0, 1]$. Note that this is just one example. There are (infinitely) many more.

For each n, define $f_n \colon [0, 1] \to \mathbb{R}$ by $f_n(x) = nx$. This gives the sequence
$f_1 \colon x \to x$
$f_2 \colon x \to 2x$
$f_3 \colon x \to 3x$
$\vdots$

Solution to Exercise 1.3

A point in Cantor space $\mathbf{C}$, is an infinite sequence of zeros and ones. So a *sequence* in Cantor space is an ordered list of sequences (in other words, a 'sequence of sequences'). Therefore the following is an example of a sequence in Cantor space. Note that this is just an example. There are (infinitely) many more:
$(0, 0, 1, 1, 0, 0, 1, \ldots)$
$(0, 1, 0, 1, 0, 1, 0, \ldots)$
$(0, 0, 0, 1, 1, 1, 0, \ldots)$
$\vdots$

Solution to Exercise 1.4

An example is the sequence $\left(\frac{1}{3^n}\right)$ (since $\frac{1}{3^n} \in K_C$ for all n and this sequence is $d^{(1)}_{K_C}$-convergent to 0).

Solution to Exercise 1.5

Assume that we have chosen the first two terms a_{N_1} and a_{N_2} as described in the proof of Lemma 1.1.

Let $P(m)$ be the proposition that we can find m terms $a_{N_1}, \ldots, a_{N_m}$ in (a_n) with $a_{N_j} \geq a_{N_i}$ for all $j \geq i$. Then $P(2)$ is true, so we have a basis for the induction. Assume that $P(k)$ is true. We know that a_{N_k} is not a peak term of the sequence (a_n) and so there must be a term a_l with $l > k$ and $a_l \geq a_k$. Set $a_{N_{k+1}} = a_l$. Then, since $a_l \geq a_k$, by the induction hypothesis we have that $a_{N_1}, \ldots, a_{N_{k+1}}$ are $k+1$ terms with $a_{N_j} \geq a_{N_i}$ for all $j \geq i$ which is $P(k+1)$. Hence, by mathematical induction, we have the result.

Solution to Exercise 2.1

Let $A \subseteq X$ be a sequentially compact subset of a metric space (X, d). To show that A is d-closed, we need to show that any d-convergent sequence with terms in A has a limit in A. Let (a_n) be a d-convergent sequence in A, with limit a. Since A is sequentially compact, (a_n) has a d_A-convergent subsequence, (a_{n_i}), with a limit, a', in A. By Book D, Chapter 15, Theorem 1.5, we have that (a_{n_i}) is d-convergent to a', and so by Book D, Chapter 16, Lemma 1.13 we have $a' = a$ and $a \in A$. Therefore A is d-closed.

Solution to Exercise 2.2

Since A is not bounded, we cannot find a real number M such that $|x| \leq M$ for each $x \in A$. Therefore for each $n \in \mathbb{N}$ we can find an element $a_n \in A$ such that $|a_n| \geq n$. This gives a sequence (a_n). Assume (for a contradiction) that we can find a convergent subsequence (a_{n_i}) with (a_{n_i}) converging to a limit L. In other words, for each $\varepsilon > 0$, there is an $N \in \mathbb{N}$ for which $d(a_{n_i}, L) < \varepsilon$ for all $i > N$. Set $\varepsilon = 1$.

See the second remark after Book D, Chapter 14, Definition 3.2.

There is nothing special about the choice of $\varepsilon = 1$ here. We just need to fix the size of a 'gap' between the possible limit and the 'tail' of the unbounded sequence.

Now $n_i \geq i$ (since (a_{n_i}) is a subsequence of (a_n)), and so for each i we have $|a_{n_i}| \geq n_i \geq i$. Pick $M > |L| + 1$. Then $|a_{n_i}| > |L| + 1$ for all $i \geq M$, and so $||a_{n_i}| - |L|| > 1$ for all $i \geq M$. Therefore, for all $i \geq M$ we have $d^{(1)}(a_{n_i}, L) = |a_{n_i} - L| \geq ||a_{n_i}| - |L|| > 1$ (by the Reverse Triangle Inequality). Therefore it is not possible to find any $N \in \mathbb{N}$ for which $d(a_{n_i}, L) < 1$ for all $i > N$, which contradicts that fact that (a_{n_i}) converges to the limit L.

Solution to Exercise 2.3

The middle-third Cantor set $K_C = \bigcap_{n=1}^{\infty} K_n$ where:

$$K_0 = [0,1]$$
$$K_1 = \left[0, \tfrac{1}{3}\right] \cup \left[\tfrac{2}{3}, 1\right]$$
$$K_2 = \left[0, \tfrac{1}{9}\right] \cup \left[\tfrac{2}{9}, \tfrac{1}{3}\right] \cup \left[\tfrac{2}{3}, \tfrac{7}{9}\right] \cup \left[\tfrac{8}{9}, 1\right]$$
$$\vdots$$

with K_n formed by removing the open middle third of each interval in K_{n-1}. Consider K_i for any i. Then K_i is a finite union of closed intervals and so, by the paragraph below Book D, Chapter 16, Proposition 1.14, K_i is closed for any i. Now $K_C = \bigcap_{n=1}^{\infty} K_n$ and so is the infinite intersection of a collection of closed sets, and so is closed by Proposition 1.17 of Chapter 16. Also $K_C \subseteq [0,1]$, and so $|c| \leq 1$ for all $c \in K_C$, and K_C is bounded. Therefore by Theorem 1.2 K_C is sequentially compact.

Solution to Exercise 2.4

Let (a_n) be a sequence in X. As (a_n) contains infinitely many terms, then at least one element of X, say x, must appear infinitely many times in the sequence. Therefore we must have $a_{n_i} = x$ for infinitely many n_i, giving a constant subsequence which is d-convergent to x (by Book D, Chapter 14, Worked Exercise 3.5).

Solution to Exercise 2.5

Let (f_n) be the sequence of functions in A given by

$$f_n(x) = n \quad \text{for each } x \in [0,1].$$

For each $x \in [0,1]$, the sequence $(f_n(x))$ diverges, and so has no pointwise limit; moreover, any subsequence of $(f_n(x))$ also diverges. Lemma 4.5 of Chapter 15 implies that there is no uniformly convergent subsequence. But, by Theorem 5.8 of Chapter 15, uniform convergence is the same as $d_{\max}$-convergence in $C[0,1]$. It follows that there are no $d_{\max}$-convergent subsequences of (f_n). Hence A is not sequentially compact.

Solution to Exercise 2.6

Let (X,d) and (Y,e) be metric spaces with X being sequentially compact. Let $f\colon X \to Y$ be a (d,e)-continuous function and let (a_n) be a sequence in the image $f(X)$. We need to find an $e_{f(X)}$-convergent subsequence. Since (a_n) is in $f(X)$, we can find a sequence (x_n) such that $f(x_n) = a_n$ for each n. Then (x_n), being a sequence in the sequentially compact space X, has a convergent subsequence (x_{n_i}), converging to a limit, say x, in X. Then, since f is continuous, we must have that $\big(f(x_{n_i})\big)$ converges to $f(x) \in f(X)$, and so $\big(f(x_{n_i})\big) = (a_{n_i})$ is a convergent subsequence of $(f(x_n)) = (a_n)$.

Solution to Exercise 2.7

Let M be such that $d^{(n)}(\mathbf{x}, \mathbf{y}) \leq M$ for all $\mathbf{x}, \mathbf{y} \in A$. Let $\mathbf{a} = (a_1, \ldots, a_n)$, $\mathbf{b} = (b_1, \ldots, b_n)$ be in A. Since $d^{(n)}(\mathbf{a}, \mathbf{b}) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \cdots + (b_n - a_n)^2}$, we have $d^{(1)}(a_i, b_i) \leq d^{(n)}(\mathbf{a}, \mathbf{b})$, where $a_i = p_i(\mathbf{a})$ and $b_i = p_i(\mathbf{b})$. Therefore $d^{(1)}\big(p_i(\mathbf{a}), p_i(\mathbf{b})\big) \leq d^{(n)}(\mathbf{a}, \mathbf{b}) \leq M$ and so $p_i(A)$ is $d^{(1)}$-bounded.

Solution to Exercise 2.8

Assume that A is bounded. Clearly the result holds for the empty set, so we may assume that A is non-empty and bounded. Then there is an $M > 0$ such that $d(x, y) \leq M$ for all $x, y \in A$.

Fix $a \in A$. Then $d(a, x) \leq M$ for all $x \in A$ and so $d(a, x) < M + \varepsilon$ for all $x \in A$ and for any $\varepsilon > 0$. Therefore $A \subseteq B_d(a, M + \varepsilon)$ for each $\varepsilon > 0$ and so $A \subseteq B_d(a, r)$, where $r = M + \varepsilon$.

Now assume that $A \subseteq B_d(a, r)$ for some $r > 0$. Then for any x and y in A, we have $d(x, y) \leq d(x, a) + d(a, y) \leq r + r = 2r$, so taking $M = 2r$ gives us an $M > 0$ such that $d(x, y) \leq M$ for all $x, y \in A$.

Solution to Exercise 2.9

By Proposition 2.11 it is sufficient to prove that S^3 is a $d^{(4)}$-closed and bounded subset of $\mathbb{R}^4$. Now S^3 is the boundary of $B_{d^{(4)}}(\mathbf{0}, 1)$ and so is $d^{(4)}$-closed. It is clearly bounded and so is sequentially compact.

Solution to Exercise 2.10

Let p_1 and p_2 be the projection functions on $\mathbb{R}^2$. Then p_1 and p_2 are continuous and

$$f(x, y) = p_2((x, y)) \sin(\pi p_1(x, y)) + p_1((x, y)) \cos(\pi p_2(x, y)).$$

It is straightforward to use our list of basic continuous functions from $\mathbb{R}$ to $\mathbb{R}$, together with the combination and Composition Rules for continuous functions from $\mathbb{R}^2$ to $\mathbb{R}$, to verify that

$$(x, y) \mapsto p_2((x, y)) \sin(\pi p_1(x, y))$$

and

$$(x, y) \mapsto p_1((x, y)) \cos(\pi p_2(x, y))$$

are continuous on $\mathbb{R}^2$ and so, by the Sum Rule, f is continuous on $\mathbb{R}^2$.

Hence, by the Restriction Rule, f is continuous on $\{(x, y) : 0 \leq x \leq 1, \text{ and } 0 \leq y \leq 1\}$.

Also, $\{(x, y) : 0 \leq x \leq 1, \text{ and } 0 \leq y \leq 1\}$ is closed and bounded. Hence, by Theorem 2.12, it is sequentially compact.

Thus both conditions of the General Extreme Value Theorem are satisfied, and we conclude that the function f attains a minimum value on $\{(x, y) : 0 \leq x \leq 1, \text{ and } 0 \leq y \leq 1\}$.

Solution to Exercise 3.1

Only if

Let $a \in X$ and suppose that f is (d, e)-continuous at $a \in X$.

Let $\varepsilon > 0$ be given. We suppose, for a contradiction, that for each $\delta > 0$, we can find $x \in B_d(a, \delta)$ for which $e(f(x), f(a)) \geq \varepsilon$. Then, in particular, for each $n \in \mathbb{N}$, we can find $x_n \in B_d(a, \frac{1}{n})$ for which $e(f(x_n), f(a)) \geq \varepsilon$. Thus for this sequence (x_n), we have that it d-converges to a, and yet $e(f(x_n), f(a)) \geq \varepsilon > 0$ for each $n \in \mathbb{N}$, meaning that the sequence $(f(x_n))$ does not e-converge to $f(a)$ – this contradicts the assumption that f is (d, e)-continuous at a.

If

For the opposite implication, let $a \in X$ and suppose that for each $\varepsilon > 0$, there is $\delta > 0$ such that $e(f(x), f(a)) < \varepsilon$, whenever $d(x, a) < \delta$.

Let (x_n) be a sequence in X that d-converges to a. Our aim is to show that $f(x_n)$ e-converges to $f(a)$.

Let $\varepsilon > 0$ be given. Then, by our assumption, we can find $\delta > 0$ for which $e(f(x), f(a)) < \varepsilon$, whenever $d(x, a) < \delta$. However, since $d(x_n, a) \to 0$ as $n \to \infty$, we can find $N \in \mathbb{N}$ so that $d(x_n, a) < \delta$ for $n > N$ and so for $n > N$, $e(f(x_n), f(a)) < \varepsilon$. In other words, for each $\varepsilon > 0$, there is $N \in \mathbb{N}$ so that $e(f(x_n), f(a)) < \varepsilon$ for $n > N$, which means that $f(x_n)$ e-converges to $f(a)$ as $n \to \infty$, and we conclude that f is (d, e)-continuous at a, as required.

Solution to Exercise 3.2

The argument is almost exactly the same as the proof of Theorem 3.3.

We prove the result using proof by contradiction: so suppose that X is sequentially compact and there is $\varepsilon > 0$ such that for each $\delta > 0$ there are $x, y \in X$ for which $d(x, y) < \delta$ and yet $e(f(x), f(y)) \geq \varepsilon$.

Then, in particular, for each $n \in \mathbb{N}$, we can find $x_n, y_n \in X$ for which $d(x_n, y_n) < \frac{1}{n}$ and yet $e(f(x_n), f(y_n)) \geq \varepsilon$.

Now consider the sequence (x_n). Since each term lies in the d-sequentially compact space X, there is a d-convergent subsequence $(x_{n_i})_{i \in \mathbb{N}}$, with limit x, say, that lies in X. Since f is (d, e)-continuous, we know that $f(x_{n_i}) \xrightarrow{d} f(x)$ as $n \to \infty$.

Let us look at the corresponding subsequence (y_{n_i}) of the sequence (y_n). Then we have $d(x_{n_i}, y_{n_i}) < \frac{1}{n_i} \leq \frac{1}{i}$ for each $i \in \mathbb{N}$ and so, since the Triangle Inequality for d implies that

$$d(x, y_{n_i}) \leq d(x, x_{n_i}) + d(x_{n_i}, y_{n_i}) \leq d(x, x_{n_i}) + \frac{1}{i},$$

we deduce, by the Squeeze Rule, that $y_{n_i} \xrightarrow{d} x$ as $i \to \infty$. Hence, since f is continuous, $f(y_{n_i}) \xrightarrow{e} f(x)$ as $i \to \infty$.

However, for each $i \in \mathbb{N}$, the definition of x_{n_i} and y_{n_i} together with the Triangle Inequality for e implies that

$$\begin{aligned} 0 < \varepsilon &\leq e(f(x_{n_i}), f(y_{n_i})) \\ &\leq e(f(x_{n_i}), f(x)) + e(f(x), f(y_{n_i})) \to 0 \text{ as } i \to \infty. \end{aligned}$$

But this is impossible, and we conclude that the theorem holds.

Solution to Exercise 3.3

Suppose that $f\colon X \to Y$ is uniformly continuous. We must show that for each $a \in X$, f is continuous at a, so let $a \in X$ be given.

We use the ε-δ characterisation of continuity given in Theorem 3.2: f is (d, e)-continuous at $a \in X$ if, and only if, for each $\varepsilon > 0$, there is $\delta > 0$ so that $e(f(x), f(a)) < \varepsilon$, whenever $d(x, a) < \delta$.

Fix $\varepsilon > 0$, then since f is uniformly continuous, there is $\delta > 0$ so that for all $x, y \in X$, $d(x, y) < \delta$ implies that $e(f(x), f(y)) < \varepsilon$. In particular, by taking $y = a$, we have the required result.

Solution to Exercise 3.4

The function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is an example of such a function. It is certainly continuous and for any $x, y \in \mathbb{R}$, we calculate

$$|f(x) - f(y)| = |x - y| \times |x + y|.$$

So if $\varepsilon > 0$ is given, then for any $\delta > 0$, taking $y = x + \frac{1}{2}\delta$ gives $|x - y| = \frac{1}{2}\delta < \delta$ and

$$|f(x) - f(y)| = \tfrac{1}{2}\delta(|2x + \tfrac{1}{2}\delta|).$$

Thus, in order for $|f(x) - f(y)| < \varepsilon$, the choice of δ depends on both x and ε; no single value of $\delta > 0$ works for all possible choices of x. In particular, given $\varepsilon > 0$ and $\delta > 0$, take $x = \frac{\varepsilon}{\delta}$ and $y = x + \frac{1}{2}\delta$, then $|x - y| < \delta$ and

$$|f(x) - f(y)| = \tfrac{1}{2}\delta\left(\left|2\frac{\varepsilon}{\delta} + \tfrac{1}{2}\delta\right|\right) \geq \varepsilon + \left(\tfrac{1}{2}\delta\right)^2 > \varepsilon.$$

Hence f is not uniformly continuous.

Solution to Exercise 3.5

For each $x \in \mathbb{R}$ and $f \in A_g$, $|f(x)| \leq |g(x)|\,(< \infty)$, and so by setting $M_x = |g(x)|$ we see that A_g is pointwise bounded.

Solution to Exercise 3.6

For $x \in [0, 1]$ and each $M > 0$, by considering the constant function $f\colon [0, 1] \to \mathbb{R}$ given by $f(x) = M + 1$, we find a function in A for which $|f(x)| = M + 1 > M$. Hence it is not possible to find a pointwise bound for A at x.

Solution to Exercise 3.7

Fix $x \in X$ and consider $f(x)$ for $f \in A$. Since f has Lipschitz constant L, we can use the Triangle Inequality (for the modulus function) to estimate

$$\begin{aligned} |f(x)| &= |(f(x) - f(a)) + f(a)| \\ &\leq Ld(x,a) + |f(a)| \\ &= Ld(x,a) + |b|, \end{aligned}$$

and this bound is independent of the function f. Hence the set A is pointwise bounded, as required.

Solution to Exercise 3.8

Let $x \in X$ and suppose $M \geq 0$. Then by considering $f\colon X \to \mathbb{R}$ given by $f(x) = M + 1$, which is certainly Lipschitz with Lipschitz constant L and hence in A, we see that $|f(x)| = M + 1 > M$, and so no pointwise bound is possible.

Solution to Exercise 3.9

If A were not pointwise bounded, then there would be $x \in [0,1]$ such that $\sup\{|f(x)| : f \in A\} = \infty$. But this means that we can find a sequence of functions (f_n) for which for all $l \in \mathbb{R}$, there is $N \in \mathbb{N}$ with $|f_n(x) - l| > 1$ for $n > N$, meaning that there could not possibly be a subsequence of (f_n) that converges to a function from $\mathbb{R}$ to $\mathbb{R}$, let alone a continuous one, and this contradicts the sequential compactness of A.

Solution to Exercise 3.10

Suppose that $B \subseteq A$. Then for each $\varepsilon > 0$, there is $\delta > 0$ so that for each $f \in A$, if $x, y \in X$ with $d(x,y) < \delta$, then $e(f(x), f(y)) < \varepsilon$. In particular, it is certainly the case that, since $B \subseteq A$, for each $f \in B$, if $x, y \in X$ with $d(x,y) < \delta$, then $e(f(x), f(y)) < \varepsilon$. Hence B is an equicontinuous set of functions, as required.

Solution to Exercise 3.11

Let $f \in A$.

Since A is equicontinuous, we know that for each $\varepsilon > 0$, there is $\delta > 0$ so that for $g \in A$, if $x, y \in X$ with $d(x,y) < \delta$, then $e(g(x), g(y)) < \varepsilon$. In particular, taking $g = f$, we have if $x, y \in X$ with $d(x,y) < \delta$, then $e(f(x), f(y)) < \varepsilon$. But this is exactly the definition of uniform continuity, and so f is uniformly continuous on X.

Finally, from Lemma 3.6 we conclude that f is continuous on X.

Solution to Exercise 3.12

(a) To show that A is sequentially compact, Theorem 3.12 tells us that it is sufficient to show that:

1. A is closed
2. A is a pointwise bounded family of functions
3. A is an equicontinuous family of functions.

A is closed

Let (f_n) be a sequence of functions in A that converges to $f \in C[0,1]$ for $d_{\max}$. We need to show that $f \in A$.

Since $f_n(0) = 0$ for each $n \in \mathbb{N}$,

$$|f(0)| = |f(0) - f_n(0)| \leq d_{\max}(f_n, f).$$

But $d_{\max}(f_n, f)$ is a null sequence, and so $f(0) = 0$.

For $x, y \in [0,1]$, the Triangle Inequality (for modulus) and the fact that $f_n \in A$ implies that, for each $n \in \mathbb{N}$,

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &\leq |f_n(x) - f_n(y)| + 2d_{\max}(f_n, f) \\ &\leq \sqrt{|x-y|} + 2d_{\max}(f_n, f). \end{aligned}$$

Since $d_{\max}(f_n, f)$ is a null sequence, it follows that

$$|f(x) - f(y)| \leq \sqrt{|x-y|} \quad \text{for each } x, y \in [0,1].$$

Hence $f \in A$ and so A is closed.

A is a pointwise bounded family of functions

Let $x \in [0,1]$. Then for each $f \in A$,

$$|f(x)| = |f(x) - f(0)| \leq \sqrt{|x-0|} = \sqrt{x} \leq 1.$$

This bound does not depend on the choice of f, and so A is a pointwise bounded family of functions.

A is an equicontinuous family of functions

Let $\varepsilon > 0$. If $x, y \in [0,1]$ and $|x - y| < \delta$ then, for each $f \in A$,

$$|f(x) - f(y)| \leq \sqrt{|x-y|} < \sqrt{\delta}.$$

In particular, if $\delta = \varepsilon^2$, then for each $f \in A$ and each $x, y \in [0,1]$, $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in [0,1]$ satisfies $|x - y| < \delta$. Thus A is an equicontinuous family of functions.

The hypotheses of Theorem 3.12 are satisfied and so A is $d_{\max}$-sequentially compact.

(b) The result of Book D, Chapter 15, Exercise 5.3 tells us that F is $(d_{\max}, d^{(1)})$-continuous and so, since F is a continuous function defined on a sequentially compact set, the General Extreme Value Theorem (Theorem 2.13) implies that F is bounded and attains its bounds on A. In particular, F attains a minimum on A.

Solution to Exercise 4.1

Let $p \in (a, b)$. Then $a < p < b$. Let $r = \min\{|p - a|, |b - p|\}$, then $B_d(p, r) \subseteq (a, b)$.

Solution to Exercise 4.2

One such subcollection is

$$\{(\tfrac{1}{10}(i-1), \tfrac{1}{10}(i+1)) : i = 0, 1, \ldots, 10\}.$$

Solution to Exercise 4.3

(a) If $0 < x \leq 1$, then there is a rational number $q > x$ and so $x \in (0, q)$. Also, $0 \in (-\frac{1}{100}, \frac{1}{100})$.

Hence

$$[0, 1] \subseteq \bigcup_{U \in \mathcal{S}_2} U.$$

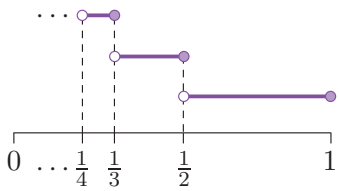
Each interval $(0, q)$ is open, and so is the interval $(-\frac{1}{100}, \frac{1}{100})$. Thus $\mathcal{S}_2$ is an open cover of $[0, 1]$.

(b) One such subcollection is the pair of sets

$$\{(-\tfrac{1}{100}, \tfrac{1}{100}), (0, q)\},$$

where q is any fixed rational number strictly larger than 1.

Solution to Exercise 4.4



One such collection is

$$\{\{0\}, (\tfrac{1}{2}, 1], (\tfrac{1}{3}, \tfrac{1}{2}], (\tfrac{1}{4}, \tfrac{1}{3}], \ldots, (\tfrac{1}{n+1}, \tfrac{1}{n}], \ldots\}.$$

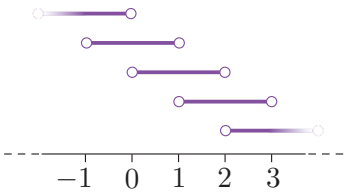
If we omit even one set, the remaining collection is not a cover of $[0, 1]$. Indeed, whatever set we omit uncovers all the points of that set.

Another possibility is given by

$$\{\{x\} : x \in [0, 1]\}.$$

This is an uncountable cover of $[0, 1]$. If we omit just a single set $\{a\}$ in this cover, then the point $a \in [0, 1]$ is no longer covered.

Solution to Exercise 4.5



There are many open covers with this property. One such is

$$\mathcal{S} = \{(-1, 1), (0, 2), \ldots, (n-2, n), \ldots\}.$$

If $x \in [0, \infty)$, then x belongs to an interval of the form $(n-2, n)$, for some $n \in \mathbb{N}$. So $\mathcal{S}$ is a cover of $[0, \infty)$. Also, the intervals in $\mathcal{S}$ are all open. So $\mathcal{S}$ is an open cover of $[0, \infty)$. But if

$$\{(n_1 - 2, n_1), \ldots, (n_k - 2, n_k)\}$$

is a finite subcollection of $\mathcal{S}$, then it does not cover any point x for which

$$x \geq \max\{n_1, n_2, \ldots, n_k\}.$$

Solution to Exercise 4.6

Let $\mathcal{S}$ be an arbitrary d-open cover of $\emptyset$. We must find a finite subcover of $\emptyset$ by sets from $\mathcal{S}$. The empty cover $\{\ \}$ is a finite subcover of $\mathcal{S}$ that covers $\emptyset$. We conclude that $\emptyset$ is compact.

Solution to Exercise 4.7

Let (X, d) be a metric space, and let A be a finite subset of X. If $A = \emptyset$, then A is compact by the result of Exercise 4.6. So suppose instead that $A = \{a_1, a_2, \ldots, a_n\} \neq \emptyset$. If $\mathcal{S}$ is an arbitrary d-open cover of A, then, for each $i = 1, 2, \ldots, n$, there is a d-open set $U_i \in \mathcal{S}$ such that $a_i \in U_i$. Hence $\mathcal{R} = \{U_i : i = 1, 2, \ldots, n\}$ is a finite subcover of A from $\mathcal{S}$. We conclude that A is compact.

Solution to Exercise 4.8

We must find an open cover of $\mathbb{R}$ that does not have a finite subcover.

An example of such a cover is given by

$$\mathcal{S} = \{(-n, n) \colon n \in \mathbb{N}\};$$

for if $\mathcal{R}$ is a finite subset of $\mathcal{S}$, then

$$\bigcup_{U \in \mathcal{R}} U = (-N, N)$$

for some $N \in \mathbb{N}$ and does not cover $\mathbb{R}$.

Solution to Exercise 5.1

Let (X, d) be a sequentially compact metric space. Lemma 5.2 implies that there is a finite ε-net $\{x_1, x_2, \ldots, x_N\}$, for some $\varepsilon > 0$. Let

$$M = \max\{d(x_i, x_j) \colon i = 1, 2, \ldots, N \text{ and } j = 1, 2, \ldots, N\}.$$

Now, by the definition of an ε-net, if $x, y \in X$ there are x_i with $d(x, x_i) < \varepsilon$, and x_j with $d(y, x_j) < \varepsilon$, and so

$$d(x, y) \leq d(x, x_i) + d(x_i, x_j) + d(x_j, y) < 2\varepsilon + M.$$

Hence X is bounded.

Solution to Exercise 5.2

(M1) Let $a, b \in X$. Then $d_{\min}(a, b) = \min\{d(a, b), 1\} \geq 0$, since d is a metric. Also since $d(a, a) = 0$ for each $a \in X$, we have $d_{\min}(a, a) = 0$ for each $x \in X$. And if $d_{\min}(a, b) = 0$, then $d(a, b) = 0$ and so $a = b$.

Thus $d_{\min}$ satisfies (M1).

(M2) Let $a, b \in X$. If $d(a, b) \leq 1$ then $d(b, a) \leq 1$ and $d_{\min}(a, b) = d_{\min}(b, a) = d(a, b)$. Also if $d(a, b) > 1$ then $d(b, a) > 1$ and we have $d_{\min}(a, b) = d_{\min}(b, a) = 1$.

Thus, for each $a, b \in X$, (M2) holds.

(M3) Let $a, b, c \in X$. We need to show that $d_{\min}(a, c) \leq d_{\min}(a, b) + d_{\min}(b, c)$. Note that if either $d(a, b)$ or $d(b, c)$ is greater than 1 then $d_{\min}(a, b) + d_{\min}(b, c) \geq 1 \geq d_{\min}(a, c)$ (as $d_{\min}(x, y) \leq 1$ for any x and y). Therefore we may assume that both $d(a, b)$ and $d(b, c)$ are strictly less than 1. In this case, $d_{\min}(a, b) + d_{\min}(b, c) = d(a, b) + d(b, c) \geq d(a, c) \geq d_{\min}(a, c)$ and so (M3) holds.

Solution to Exercise 5.3

First assume that (a_n) is d-convergent to a limit, say $L \in X$. Then, by definition, $(d(a_n, L))$ is a (real) null sequence. To show that (a_n) is $d_{\min}$-convergent, we need to prove that $(d_{\min}(a_n, L))$ is a null sequence. However, $d_{\min}(a, b) \leq d(a, b)$ for any $a, b \in X$ and so, by the Squeeze Rule, (Book D, Chapter 13, Lemma 1.15) the sequence (a_n) is $d_{\min}$-convergent to L.

Now assume that (a_n) is $d_{\min}$-convergent to a limit, say L. Then $(d_{\min}(a_n, L))$ is a null sequence, and so for each $\varepsilon > 0$, there is an $N \in \mathbb{N}$ such that $d_{\min}(a_n, L) < \varepsilon$ whenever $n > N$. We need to prove that $(d(a_n, L))$ is a (real) null sequence. In other words, we need to show that for each $\varepsilon > 0$, there is an $N \in \mathbb{N}$ such that $d(a_n, L) < \varepsilon$ whenever $n > N$.

Let $\varepsilon > 0$ be given and set $\varepsilon' = \min\{\varepsilon, \frac{1}{2}\}$. Then, since (a_n) is $d_{\min}$-convergent, we can find an $N \in \mathbb{N}$ such that $d_{\min}(a_n, L) < \varepsilon'$ whenever $n > N$. Since $\varepsilon' < 1$, we have that $d_{\min}(a_n, L) < 1$ whenever $n > N$, and so we must have $d_{\min}(a_n, L) = d(a_n, L)$ whenever $n > N$. Therefore we have $d(a_n, L) < \varepsilon'$ whenever $n > N$. Hence $(d(a_n, L))$ is a null sequence and (a_n) is d-convergent with limit L.

Solution to Exercise 5.4

We need to show that there is $M > 0$ such that $d_{\min}(x, y) \leq M$ for all $x, y \in A$. Taking $M = 2$ gives the result. (Any number greater than 1 will suffice.)

Solution to Exercise 5.5

We must show that for each $\varepsilon > 0$, there is a finite ε-net. Let $\varepsilon > 0$. If $\varepsilon > (b - a)$, then $\{a\}$ is an ε-net for $[a, b]$. If $0 < \varepsilon \leq (a - b)$, then, for example,

$$\{a + \tfrac{1}{2}n\varepsilon : n \in \mathbb{N} \text{ and } 0 \leq n \leq \tfrac{2(b-a)}{\varepsilon}\}$$

is a finite ε-net for $[a, b]$ with the Euclidean metric.

Solution to Exercise 5.6

Let (X, d) be a compact metric space, and let $\varepsilon > 0$. We seek a finite ε-net for X. The collection

$$\{B_d(x, \varepsilon) : x \in X\}$$

is an open cover of X. Since X is compact, the cover has a finite subcover

$$\{B_d(x_1, \varepsilon), B_d(x_2, \varepsilon), \ldots, B_d(x_n, \varepsilon)\}.$$

So $\{x_1, x_2, \ldots, x_n\}$ is a finite ε-net. Since $\varepsilon > 0$ is arbitrary, X is totally bounded.

Solution to Exercise 6.1

Let $\mathcal{S}$ be an open cover of $\bigcup_{m=1}^{n} K_m$. Then $\mathcal{S}$ is an open cover of K_m for $m = 1, 2, \ldots, n$ and so for each $m = 1, 2, \ldots, n$, we can find a finite subcover of $\mathcal{S}$, $\mathcal{S}_m$ say, such that

$$K_m \subseteq \bigcup_{U \in \mathcal{S}_m} U.$$

But then $\{U : U \in \mathcal{S}_m \text{ for some } m = 1, 2, \ldots, n\}$ is a finite collection of open sets from $\mathcal{S}$ that covers $\bigcup_{m=1}^{n} K_m$. Hence $\mathcal{S}$ has a finite subcover of $\bigcup_{m=1}^{n} K_m$ and the result follows.